Universidad del Desarrollo
Facultad de Gobierno
Centro de Investigación en Complejidad Social

CHALLENGES OF SOCIAL COHESION OF IMMIGRANTS

BY: YERKA TAMARA INDALICIA FREIRE VIDAL

Thesis submitted to the Faculty of Government of the Universidad Del Desarrollo for the academic degree of Doctor in Sciences of Social Complexity

SUPERVISOR THESIS:
Sr. JOSÉ ANTONIO MUÑOZ

July, 2023

SANTIAGO, CHILE

Declaration of Authorship

Dedication

Dedicated to my loved ones, to anyone to whom my work may be useful, and to Icochan.

ACKNOWLEDGMENTS

At this moment of doctoral completion, I wish to express my deep gratitude to all those people who have been instrumental in making this research project a reality and who have contributed significantly to the culmination of this doctoral thesis.

First of all, I would like to thank my professor and guide during my doctoral years, Eduardo Graells Garrido, whose expert guidance, constant support and wise suggestions have been invaluable for the development of this work and my training in general. Thank you for being a constant source of inspiration and for providing me with the necessary guidance to overcome every obstacle and challenge.

Likewise, I thank my thesis tutor, José Antonio Muñoz, who welcomed me and motivated me at times when the battle seemed lost, and with his wisdom and warmth he knew how to direct my gaze to the goal that was closer than I could see.

I sincerely thank the members of my thesis committee, Carlos Rodríguez Sickert and Werner Creixell, for their valuable contributions and for dedicating their time and effort to review and evaluate this work. Their comments and suggestions have significantly enriched the content of this thesis.

I cannot fail to mention my fellow doctoral students and friends who have shared with me their knowledge, experiences and friendship throughout this arduous journey; especially Gabriela Valdebenito, María Paz Martínez, Sebastián Diaz, Mauricio Aspé, Carina Sturla, Víctor Márquez, Ignacio Toledo, Jorge Urrutia, Jorge Castillo, Víctor Landaeta, Gabriela Fajardo, Oriana Figueroa, Paula Vásquez e Ignacio Pérez. Undoubtedly, their company, conversations and discussions have been fundamental for my academic and personal growth.

Special thanks go to my parents, María Vidal and Jorge Freire, my partner, Daniel Suárez, my cousin, María Isabel Godoy, my family and friends, whose unconditional support and constant encouragement gave me the strength to overcome difficult moments and keep me focused on my goals. Their words of encouragement and love have been my greatest motivation.

Finally, I would like to express my gratitude to all those who, in one way or another, have contributed to this research and have left an imprint on my life. To my professors, classmates, and all the people who work or worked at the University while I was there, I thank you for your good vibes and help in the most varied topics one can imagine.

This achievement would not have been possible without the support and contribution of each of the people mentioned above and many others who, in some way, have been part of this journey. To all of you, my deepest thanks.

Resumen

La migración es un fenómeno que ha adquirido gran relevancia en la actualidad, y que puede generar diversas actitudes en las sociedades receptoras. Muchos países han puesto foco en entender y reducir las actitudes discriminatorias y xenofóbicas hacia los inmigrantes, dado que dichas actitudes imposibilitan la cohesión social y la sana convivencia, además de, fomentar la marginalización de los inmigrantes. Las fuentes existentes de datos para analizar actitudes hacia la inmigración, en su mayoría encuestas y/o diseños experimentales de casos de estudio específicos, entregan valiosa información; sin embargo, imponen algunas dificultades ya que requieren de muchos recursos para implementarse y analizarse; en términos de dinero, tiempo y personas. Además, debido al espaciamiento temporal con el que se obtienen estos datos, se dificulta el acceso a la dinámica del fenómeno, y por tanto, sólo es posible obtener una imagen parcial de lo que sucede en una sociedad que cambia rápida y continuamente. En la actualidad, las redes sociales brindan una oportunidad para complementar y superar algunas de las limitaciones de los datos recopilados a través de estos medios tradicionales. Por ejemplo, Twitter no sólo sirve como un espacio público para el intercambio de opiniones e ideas sobre diversos temas sociales, sino que también influye en las opiniones de sus usuarios. En este trabajo de tesis, se presentan tres estudios que buscan inferir y analizar las actitudes hacia la inmigración, tomando como caso de estudio a la sociedad chilena, y utilizando datos de Twitter como fuente de información. El primer estudio presenta una metodología basada en el análisis de tópicos sobre los datos de Twitter para medir, clasificar y caracterizar actitudes. El segundo y tercer estudio, en cambio, usa un clasificador XGBoost para inferir actitudes hacia la inmigración. En particular, el segundo estudio muestra un análisis comparativo pre- y post-pandemia de COVID-19 para testear una de las hipótesis del Sistema Inmunológico Conductual (SIC); que indica que en un contexto pandémico deberían acentuarse las actitudes discriminatorias y xenofóbicas hacia los inmigrantes. Finalmente, el tercer estudio, expone un marco metodológico para el análisis de dichas actitudes desde una perspectiva que caracteriza el contenido, las dimensiones psicolingüísticas, y la dinámica asociada a dichas actitudes. En terminos generales, encontramos que las actitudes hacia la inmigración parecen estar influenciadas por eventos noticiosos relacionados a la migración. Además, en el uso del lenguaje, los usuarios de actitudes positivas revelan mayor empatía, mientras que los de actitudes negativas muestran una mayor percepción de amenaza; siendo consistente con las teorías sociales que explican las diferentes actitudes. Finalmente, encontramos que los usuarios de actitudes negativas son más vociferantes, incluso como un efecto de la pandemia de Covid-19, pese a que no encontramos evidencia robusta que apoye la hipótesis del SIC. Nuestro trabajo presenta novedosas metodologías para estudiar actitudes hacia la inmigración usando datos de redes sociales, y aporta valiosa información para el contexto migratorio chileno. Nuestros resultados podrían apoyar en el diseño de políticas públicas y sociales adecuadas que permitan una integración efectiva y pacífica de los inmigrantes.

Abstract

Migration is a phenomenon that has acquired great relevance nowadays, and that can generate diverse attitudes in receiving societies. Many countries have focused on understanding and reducing discriminatory and xenophobic attitudes towards immigrants, since such attitudes make social cohesion and healthy coexistence impossible, as well as fostering the marginalization of immigrants. Existing data sources for analyzing attitudes towards immigration, mostly surveys and/or experimental designs of specific case studies, provide valuable information; however, they impose some difficulties as they require a lot of resources to implement and analyze; in terms of money, time and people. In addition, due to the temporal spacing with which these data are obtained, it is difficult to access the dynamics of the phenomenon, and therefore, it is only possible to obtain a partial picture of what is happening in a society that is changing rapidly and continuously. Today, social networks provide an opportunity to complement and overcome some of the limitations of the data collected through these traditional means. For example, Twitter not only serves as a public space for the exchange of opinions and ideas on various social issues, but also influences the opinions of its users. In this thesis work, three studies are presented that seek to infer and analyze attitudes towards immigration, taking Chilean society as a case study, and using Twitter data as a source of information. The first study presents a methodology based on topical analysis of Twitter data to measure, classify and characterize attitudes. The second and third studies, on the other hand, use an XGBoost classifier to infer attitudes towards immigration. In particular, the second study shows a comparative pre- and post-pandemic analysis of COVID-19 to test one of the Behavioral Immune System (BIS) hypotheses; which indicates that in a pandemic context discriminatory and xenophobic attitudes towards immigrants should be accentuated. Finally, the third study presents a methodological framework for the analysis of these attitudes from a perspective that characterizes the content, the psycholinguistic dimensions, and the dynamics associated with these attitudes. In general terms, we found that attitudes towards immigration seem to be influenced by news events related to migration. Furthermore, in the use of language, users with positive attitudes reveal greater empathy, while those with negative attitudes show a greater perception of threat, consistent with the social theories that explain the different attitudes. Finally, we find that negative attitude users are more vociferous, even as an effect of the Covid-19 pandemic, even though we find no robust evidence to support the BIS hypothesis. Our work presents novel methodologies to study attitudes towards immigration using social network data, and provides valuable information for the Chilean migration context. Our results could support the design of appropriate public and social policies that allow for an effective and peaceful integration of immigrants.

Contents

List of Figures

List of Tables

List of abbreviations

XGBoost Extreme Gradient Boosting

SIC Sistema Inmunológico Conductual

BIS Behavioral Immune System

URL Uniform Resource Locator

TS-NMF Topic-Supervised Non-Negative Matrix Factorization

LOWESS Locally Weighted Scatterplot Smoothing

LIWC Linguistic Inquiry and Word Count

UN United Nations

DESA Department of Economic and Social Affairs

U.S. United States

UK United Kingdom

ESS European Social Survey

API Application Program Interface

TF-IDF Term Frequency – Inverse Document Frequency

SCC Strongly Connected Component

HIV Human Immunodeficiency Virus

RTs Retweets

ICT Intergroup Contact Theory

ITT Integrated Threat Theory

RQ Research Question

Chapter 1

INTRODUCTION

Human migration is the phenomenon by which people move from one place to another, either within a country (internal migration) or from one country to another (international migration). People may migrate for a variety of reasons, such as seeking better economic opportunities, escaping armed conflict or political persecution, reuniting with family members, seeking better living conditions, among others. This type of migratory displacement has been present in all periods of worldwide history, and appears to be a growing effect of current globalization. In 2020, the number of international migrants reached 281 million worldwide, equivalent to 3.6% of the total world population (according to the report: "International Migration 2020 Highlights”, by the Population Division of the UN, DESA). Although migration is a universal human right [148] , it is far from being a phenomenon free of associated social problems. Many countries face migration processes that have a variety of effects on both the emigrating and receiving populations.

While migration can have positive consequences, such as relieving pressure on resources in countries of origin and providing a flow of remittances to families and communities left behind, as well as contributing to the economy of destination countries through labor and tax payments, what is of concern are the negative impacts that it entails, such as the loss of skilled labor in countries of origin, while in destination countries there can be tensions related to the integration of migrants and competition for jobs and resources, among others.

Countless studies have been developed that seek to understand the determining factors in the formation of such effects.

On the one hand; from the field of economics, some studies have linked immigration to the improvement or worsening of the economy. However, the results are not conclusive. In 1997 George J. Borjas et al. argued that Immigration does not have a consistent, discernible effect on area economic outcomes and other regional factors dominate the ups and downs of area economies [18]. Moreover, David Card (2001) finds that “immigrant inflows over the 1980s reduced wages and employment rates of low-skilled natives in traditional gateway cities like Miami and Los Angeles by 1–3 percentage points” [26]. Similarly, George J. Borjas (2003) finds that immigration in the U.S. lowers the wage of competing workers: a 10 percent increase in supply reduces wages by 3 to 4 percent [17]. More recently, Konrad B. Burchardi et al. (2020) [23] finds that immigration has a positive causal impact on innovation, measured as patents of local firms, and on economic growth, measured as real income growth of native workers, for a U.S. study.

On the other hand, from the social sciences, efforts have focused on the problems of integration and social cohesion, because the interaction between locals and immigrants is, often, more complex than the interactions between people from the same country or region; since locals and immigrants must face important differences such as: social identity, language, race, religion and traditions. On a psychological level, John W. Berry finds that in a process of acculturation, those who pursue and encourage marginalization experience more stress than those who favor integration [13]. In addition, it is common for immigrants who do not adapt well in a society to suffer the effects of social marginalization and for socioeconomic inequalities in society to be accentuated. For example, Algan et al. (2010) show that the local population has higher employment, education and income than ethnic minorities in the UK, Germany and France [2]. Eriksson (2010) finds these same disparities in Sweden [45].

Therefore, it is fundamental to know if discrimination is present, because it is possible that this will negatively affect the integration of immigrants, this would most likely further enforce inequality. Whether by their own decision or by discriminatory attitudes of the locals toward them, immigrants are often segregated in different social spheres: housing, work, education, etc. For example, isolated communities of immigrants (ghettos) are created, which although they represent some advantages for immigrants (for example, sharing customs, speaking the same language and maintaining a social identity), often maintain and accentuate the differences with the locals. The European Social Survey 2017 (ESS) shows that people with a migrant background are better integrated in inclusive countries [46]. Being an inclusive society requires, by definition, not to discriminate and therefore, that people do not have negative or hostile attitudes towards immigrants. Thus, immigration sentiment is an essential component for successful migrant integration into receiving societies.

The attitudes towards immigration are very varied, there are those who support multiculturalism and integration, as well as those who think that immigrants will bring negative effects for their society. Since, communities that have negative attitudes toward minorities (immigrants, for example) tend to hinder adaptation, with averse majorities playing a key role in the emergence of segregated minority communities [126], and these attitudes can lead to acts of violence, discrimination and abuse; it is important to understand and identify factors that enhance these attitudes. This knowledge could prevent, inform and even support public policies that help to mitigate negative attitudes, enhance positive ones, and foster social cohesion.

In order to understand the formation of attitudes towards immigration there are two main theories: “The Intergroup Contact Theory” that says that the contact between people from different groups will foster empathy and improve relationships [3] and conversely, “The Integrated Threat Theory” says that the contact between people from different groups will cause prejudices and perceptions of threat, which will worsen relationships [112, 141].

However, measuring attitudes is not yet a simple or solved problem. Traditional data sources are based on surveys and specific experimental studies, which makes them difficult to replicate in other contexts or geographic regions. Moreover, they are costly to implement and analyze, in terms of time and money, so having such data for a fine granularity of time is almost impossible. Thus, studying the dynamics of attitudes is an even greater challenge.

The main contribution of this thesis work focuses on this limitation. We propose to make use of the information that Twitter users publish in their accounts. It is very common to find reactions and attitudes through posts in social networks, where people express their ideas and opinions freely and voluntarily. Twitter is currently a platform widely used in studies of human behavior, since it provides a valuable source of data. Studies that have used Twitter have allowed to reveal socio-cultural characteristics of users or societies, including the level of integration of immigrants in a city [100], attitudes in response to triggering events, such as terrorist attacks [39], the influence of culture in personal actions [63], political polarization [63], personality traits [128], and personality differences between democrats and republicans [144]. Thus, the main objective of our research is to measure, classify and characterize attitudes towards migration using the migration debate on Twitter as a proxy.

We take Chilean society as a case study because, while immigration is not a new phenomenon in Chile, currently it has become a very contingent topic because immigrant population has drastically increased in recent years. The percentage of immigrants increased, from 0.8% in 1992 to 4,35% in 2017, and 66.7% of immigrants declare to have arrived in Chile between 2010-2017; mainly from 2016 [94]. And these rates continue to rise; it is estimated that by 2021, the immigrant population will account for 7.6% of the total population. Also, the composition of immigrants has also changed. Haiti being the most notorious case since they have a different native language (créol), their country is one of the poorest in the world, and most of them are of Afro-Haitian descent, an ethnicity that was almost non-present in Chile. In addition, even during the global pandemic of Covid-19, Chile has faced some migratory crises, mainly due to the Venezuelan exodus, resulting in a large increase in the immigrant population entering the country irregularly [93]. For this, Chileans have developed diverse perceptions regarding the number of immigrants in the country and the phenomenon itself. This makes Chile a relevant case study for attitudes toward immigration.

The present thesis work consists of three specific studies on attitudes towards migration, two of which have already been published (see appendix in chapter 5), and the third is in the final process of submission to a journal. Each of them is addressed in a separate chapter. The thesis is organized as follows; the Chapter 2 shows a study published and titled as “Characterization of Local Attitudes Toward Immigration Using Social Media”, in it we use a semi-supervised topic modeling named Temporal-Spatial Non-negative Matrix Factorization [107], to place 49K users and more than 206K tweets, into a spectrum ranging from in-favor to against immigration. We characterized both sides of the spectrum in two aspects: the emotions and lexical categories relevant for each attitude, and the discussion network structure. In Chapter 3 we present a study in progress entitled “Attitudes towards migrants in a COVID-19 context: testing a Behavioral Immune System hypothesis with Twitter data”, in which we use the XGBoost classifier [30], an ensemble machine learning algorithm, widely used for classification problems, to test a hypothesis of the Behavioral Immune System that indicates that, in a pandemic context, discriminatory and xenophobic attitudes towards migrants should be accentuated [137]. Thus, in this research we present a comparative analysis, in which we analyze the attitudes towards migration of more than 21k users and 310k tweets, in a pre- and post-pandemic period of Covid-19. And finally, in Chapter 4, we present a study published and titled as “A Framework to Understand Attitudes towards Immigration through Twitter”, in which we apply the same methodology used in the second study, and with it we define a framework that allows the classification of more than 36k users and 160k tweets into positive and negative attitudes towards immigrants, and provides a characterization of these profiles by quantitatively summarizing user content and temporal trends of the posts.

The order in which the last two case studies are presented does not correspond specifically to the order in which they were developed during the thesis work. The reason for this is explained below. Although the work presented in Chapter 4 was developed prior to the work presented in Chapter 3, we chose this order of presentation because both studies share the methodology applied to measure and classify attitudes toward migration. However, the methodology is presented in greater detail in the study that has not yet been published (Chapter 3), so we prefer to present it first. The difference in the level of detail of methodological explanation between the two studies lies in the audience for which they were developed. On the one hand, the paper in Chapter 3 is for a more social science oriented audience so we have explained more deeply the computational methods used, and on the other hand, the paper in Chapter 4 was published in a journal with an audience closer to the computational sciences, which allows us to explain in a less detailed way the methodological part.

With the results of these analyses, we examine the integration of immigrants into society from the perspective of attitudes towards migration. We contribute to the understanding of the migratory sentiment in Chile and the characteristics of users who show positive and negative attitudes, respectively; and more generally, we provide a novel methodological framework that allows us to extract new sources of information based on social networks. We believe our results help to inform the design of public policy and interventions to improve relations between groups in a country.

Chapter 2

STUDY 1. CHARACTERIZATION OF LOCAL ATTITUDES TOWARD IMMIGRATION USING SOCIAL MEDIA

Abstract

Migration is a worldwide phenomenon that may generate different reactions in the population. Attitudes vary from those that support multiculturalism and communion between locals and foreigners, to contempt and hatred toward immigrants. Since anti-immigration attitudes are often materialized in acts of violence and discrimination, it is important to identify factors that characterize these attitudes. However, doing so is expensive and impractical, as traditional methods require enormous efforts to collect data. In this study, we propose to leverage Twitter to characterize local attitudes toward immigration, with a case study on Chile, where immigrant population has drastically increased in recent years. Using semi-supervised topic modeling, we situated 49K users into a spectrum ranging from in-favor to against immigration. We characterized both sides of the spectrum in two aspects: the emotions and lexical categories relevant for each attitude, and the discussion network structure. We found that the discussion is mostly driven by Haitian immigration; that there are temporal trends in tendency and polarity of discussion; and that assortative behavior on the network differs with respect to attitude. These insights may inform policy makers on how people feel with respect to migration, with potential implications on communication of policy and the design of interventions to improve inter-group relations.

Keywords: Immigration; Public attitude; Twitter; Semi-supervised Topic model.

2.1 Introduction

Migration is a phenomenon faced by many countries, which brings a variety of effects; both in the population from which it emigrates and in the receiving population. One of the effects that worries many countries is intolerance and hostile attitudes toward immigrants. These attitudes have been the focus of many research studies, some of which are focused on individual-level psychological and socio-economic factors [24, 138], and others on the contact between immigrant population and locals [21, 91, 95]. The main methods used in these studies are based on context specific surveys, which makes replication in others societies or countries difficult. The theories that explain the type of attitudes of locals interacting with immigrants can be summarized in two: the Intergroup Contact Theory [3], and the Integrated Threat Theory [141, 112]. The former states that people support multiculturalism and integration. The latter, that people think that immigrants will bring negative effects for their society, including competition for jobs and public services, worsening of the national economy, increase in crime, and the arrival of diseases. Particularly, the attitudes explained by the threat theory can lead to acts of violence, discrimination, and abuse; thus, it is important to understand what factors enhance such attitudes.

However, measuring attitudes is costly and impractical under dynamic scenarios. The most frequent methods are surveys, which are difficult and costly to implement. In this study, we propose to make use of the information that people publish in Twitter as a proxy of their attitudes toward immigration. It is common to find reactions and attitudes through posts in these platforms, where people express their ideas and opinions voluntarily. We propose to define a spectrum of attitudes based on the two aforementioned theories, and to classify users and tweets into that spectrum. We do so with a semi-supervised topic modeling technique named Topic-Supervised Non-Negative Matrix Factorization [107]. TS-NMF works in a semi-supervised way because some users can be labeled as belonging to each extreme of the spectrum, something that we do with custom-built lexicons for each theory.

We perform a descriptive case study on the Chilean society, because Chile is one of the countries in which migration has reached unprecedented volume in recent years. The statistics show that immigrant population has increased from 0.8% in 1992 to 4.35% in 2017; and where 66.7% of immigrants declare to have arrived mainly in 2016 [94]. For this, Chileans have developed diverse perceptions regarding the number of immigrants in the country and the phenomenon itself. To measure them with our proposed method, we collected more than 206K tweets that discuss immigration in Chile, written by more than 49K users during the year 2017. After inferring user and tweets positions in the spectrum, we performed lexical and network analysis with respect to the spectrum position. In the lexical analysis, we used the "Linguistic Inquiry and Word Count" (LIWC) lexicon [119], typically employed to characterize cognitive and emotional differences in discourse [87, 36, 69]. To analyze the network structure, we estimated the polarization of the retweet and mention networks between users.

As main results, we observed that most of the discussion toward migration in Chile is targeted at Haitian migration, even though other countries have a larger share of the population. We found lexical differences in how each attitude discussed migration, and those differences were consistent with theories. For instance, social-related words were correlated with empathetic attitudes, job- and money-related words were correlated with threatening attitudes. In the network, the retweet network was polarized, in coherence as predicted by other studies regarding political discussion [34, 75]. Finally, we notice that the amount and tendency of the tweets (the latter reflects the attitude towards immigration) seems to be influenced by relevant news events on national migration issues. These results can inform public policy designers to improve inter-group relations in the country, as well as increasing the understanding of how people feel regarding an important aspect of globalization.

In summary, the contribution of this study is two-fold. We proposed a methodology to characterize local attitudes toward migration from tweets. Then, we performed a descriptive case study in Chile using this method, obtaining results that are coherent with social theory, with added depth based on the rich information that can be extracted from Twitter.

This chapter is structured as follows. Section 2.2 discusses the related work and describes the social theories that guided our analysis. Section 2.3 describes the data set we analyzed. Section 2.4 describes the methodology. Section 2.5 describes the results of applying the methodology to the data set. Section 2.6 discusses the implications of our work. Finally, Section 2.7 states our conclusions.

2.2 Related Work

Migration is a widely studied topic because there are many issues associated to this phenomenon. Some researchers have focused on studying the economic impacts related to migration [138, 83, 86, 84], others on social cohesion [140, 91, 96, 95]. Within these studies, those who have focused on integration [100] and racism/xenophobia stand out [21, 88, 122]. Our work seeks to contribute in the latter area, mainly due to the subject of our case study, Chile, a society that in a short time has faced a massive influx of immigrants. Migration in Chile has been a national issue, causing controversy in presidential elections, news, and municipal institutions. However, measuring attitudes is not a simple problem, nor a solved one. Twitter is currently a platform widely used in studies of human behavior, since it provides a valuable source of data. Studies that have used Twitter have allowed to reveal socio-cultural characteristics of users or societies, including the level of integration of immigrants in a city [100], attitudes in response to triggering events, such as terrorist attacks [39], the influence of culture in personal actions [63], political polarization [63], personality traits [128], and personality differences between democrats and republicans [144].

Given this body of research, we propose that Twitter can be used as a proxy to understand human behavior, in our case, the attitudes of Chileans regarding immigration.

2.2.1 Social Theories

The attitudes toward immigration are varied and depend on economic, socio-cultural and psychological factors. In this context, psychology and sociology have defined theories that explain the attitudes exhibited by people, who belong to different groups, when interacting with others: the Intergroup Contact Theory [3], and the Integrated Threat Theory [141, 112]. The attitudes toward immigration are a particular case explained by these theories.

Intergroup Contact Theory. Developed in the book "The Nature of Prejudice" by Gordon W. Allport [3], it postulates that prejudices are reduced when the interaction between different groups meets the following conditions: 1) groups are on equal terms; 2) they have common goals; 3) there is cooperation; and 4) there is support from formal and/or informal institutions. The theory states that intergroup contact reduces the fear and anxiety that exists when people interact with an unknown group [143], and that it promotes empathy and understanding towards the foreign group [142].

This theory has been used to ground several studies: contact between white and black people [21], heterosexuals and homosexuals [88, 89], minority religious groups [115], and locals and immigrants [91]. All these studies conclude that contact improves relationships between groups.

Integrated Threat Theory. In contrast to the Intergroup Contact Theory, the Integrated Threat Theory argues that contact between disparate groups provokes perceptions of threat and contempt [141, 112], for instance, due to competition for work and economic resources [82, 47]. Furthermore, the threat does not have to be real, it can be subjective or fictitious [96].

This theory postulates that, when the interaction conditions are not optimal, the contact between different groups will provoke conflicting and hostile relationships. The concept of "contact" is not limited only to physical contact, it can also be indirect [44], imagined [37], and electronic [4, 153].

Both theories tell us what to search when we look attitudes toward immigration: attitudes motivated by empathy, in favor of immigration; and attitudes motivated by threat, against immigration. As such, we will assume that there are two attitudes, which we label as *empathy* and *threat*.

2.3 Dataset Description

In this section we describe our data set of posts from Twitter about migration in Chile.

Twitter is a micro-blogging platform, where users publish tweets (posts) with a maximum of 280 characters. Users may follow others, to see their tweets in their own timelines. Tweets may mention other users, quote other tweets, or retweet another tweet to share it with one's audience. Users can report a screen name, a full name (which can be real or fictitious), a location (real, fictitious, or empty), and a small autobiography, among other attributes. To collect tweets that talk about immigration in Chile we used the Twitter Streaming API using system designed to

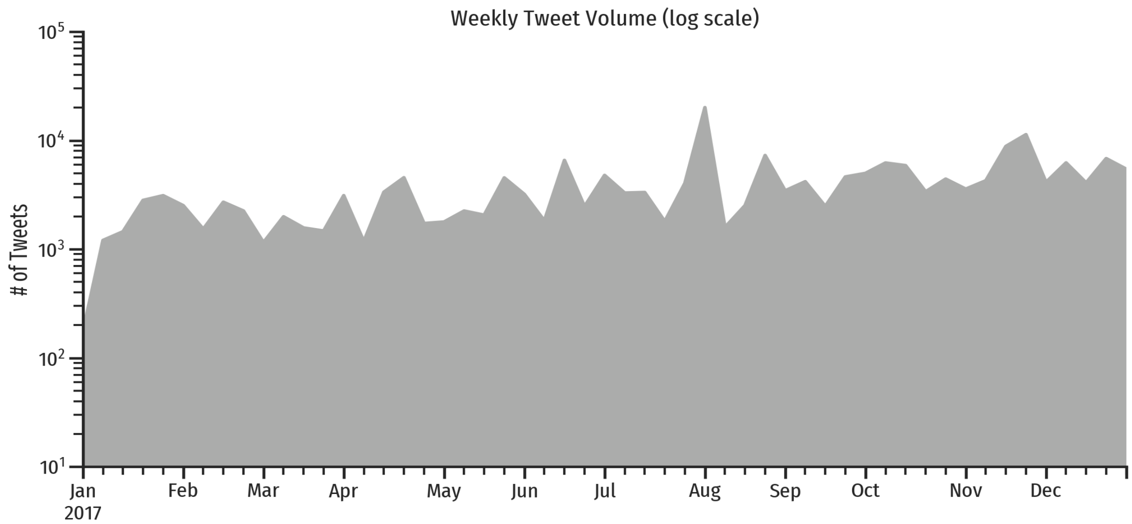


Figure 2.1: Weekly distribution of tweets about immigration in Chile, year 2017.

crawl Chilean tweets [77]. The query parameters were keywords related to immigration (*e.g.*, inmigración, inmigrante, fronteras, racismo, *etc.*), and origin countries with their respective demonyms (*e.g.*, Haití–haitianos/as, Venezuela–venezolanos/as, Perú–peruanos/as, *etc.*). Given how generic some of these keywords are, particularly regarding the context of political issues of neighbouring countries, and the presidential elections held in Chile during November and December, we performed extensive manual clean-up of the data set.

In total, our data set is comprised by 206,353 tweets that discuss immigration in Chile during 2017, written by 49,346 users. Figure 2.1 shows the weekly volume of tweets. As seen on the figure, the amount of tweets has a sligth positive trend. Two peaks draw our attention: July 31th, when the news reported a case of an Haitian citizen with Leprosy; and November 19th, when an Haitian citizen rescued a woman who fell from the ninth floor of a building.

Regarding content, Figure 2.2 shows the most frequent words, after removing stopwords and accents. One can see that words such as Haití and Haitianos

Figure 2.2: Wordcloud of most frequent words in the dataset, after removing stop-words. Color is assigned according to the following categories: words, hashtags, mentions, and URLs.

are more relevant than other countries name or demonyms, despite the fact that the largest immigrant population comes from Perú, Colombia and Venezuela [94]. Also, Santiago and Antofagasta are two frequent keywords, the two cities with the largest immigrant population [94]. Other relevant words that appear are: "gobierno" (government), "carabineros" (policemen), "Piñera" (current president, and presidential candidate in 2017) and "proyecto" (project); possibly because during the year an immigration reform was being discussed.

Table 2.1: Examples of training terms for each attitude.

Attitude	**Training Terms**
Empathy	#todossomosmigrantes, #stopxenophobia, #chilesinbarreras, #chileterecibe, #bienvenidosmigrantes, @oimchile, bienvenidos a chile, #derribandomuros, @sjmchile, ...
Threat	#vendepatria, #nomasinmigrantes, #nomasilegales, #inmigrantesilegales, inmigrantes delincuentes, inmigracion descontrolada, indeseables, ...

We explored the data set to seek for words, phrases, and hashtags that could be mapped to the *empathy* and *threat* attitudes. In *empathy* we chose terms that indicated that immigrants are welcome and will be received in equal conditions (*e.g.*, “we are all immigrants”). In *threat* we chose terms and words that showed that immigrants are not welcome and qualified them negatively (*e.g.*, “illegal immigrants”). Table 2.1 shows some examples of the the terms we associated to both attitudes. These labeled terms are not necessarily frequent, however, the methodology that we describe in the next sections allows to propagate these labels through a topic model.

2.4 Methodology

In this section we describe how to characterize users and tweets according to their attitude toward immigration. We define how to apply machine learning techniques to user profiles to derive user-attitude and term-attitude associations. Then, we define how to characterize attitudes from sentiment, lexical and network perspectives.

2.4.1 Attitudes and Topic Modeling

Topic models are a family of techniques used to discover the underlying semantic structure of a corpus by identifying and quantifying the importance of representative themes in all documents [15]. Topic models assume that each text document is generated by a set of topics which have a determined distribution. At the same time, each topic is defined by a set of words, which also have a particular distribution for each topic.

A popular topic modeling technique is Non-negative Matrix Factorization [102].

NMF works by constructing a k-rank factorization of a positive document-term matrix V into $W \times H$. Matrices W and H are estimated by minimizing the following objective function:

$$D_{NMF}(W, H) = \parallel V - W \times H \parallel_F^2, \quad W, H \geq 0, \tag{2.1}$$

where $\parallel \cdot \parallel_F$ is the Frobenius norm. In topic modeling, k, W and H have a special interpretation: k is the number of topics, W_{ij} quantifies the relevance of topic j in document i, and H_{ij} quantifies the relevance of term j in topic i.

Typical topic modeling applications select different numbers of k based on metrics such as perplexity. However, the meaning of topics is not always interpretable, as the factorization may follow latent patterns not necessarily aligned with human expectations. Based on the social theories described in Section 2.2.1, we propose to guide the learning procedure to seek for two topics: one that represents *empathy*, and another that represents *threat*. In such cases, supervised methods could be employed, however, these methods require a fully labeled data set, not available in our case. Since it is possible to map specific terms (words, phrases, hashtags, URLs, *etc.*) into these two topics, we propose to use a semi-supervised version of NMF known as Topic-Supervised NMF [107]. TS-NMF defines the minimization problem as follows:

$$D_{TS}(W, H) = \parallel V - (W \circ L)H \parallel_F^2, \quad W, H \geq 0, \tag{2.2}$$

where $\circ$ is the Haddamard product operator, and L is a supervision matrix, defined as $L_{ij} = 1$ if topic j contributes to the document i, and $L_{ij} = 0$ if the topic j does not contribute to the document i. Thus, TS-NMF allows to provide examples of documents labeled with known topics, and to restrict the latent representation of the corpus to align with the labeled examples.

In our context, we work with user profiles, *i.e.*, the concatenation of tweets by a single user is one document. As terms we consider hashtags, mentions, URLs, and n-grams with n up to four. This allows us to define how specific phrases are mapped to each topic. The user corpus is represented as a document-term matrix D weighted with TF-IDF [6], and then row-normalized with L2 norm. To label users in the supervision matrix, we construct a list of seed terms for each theory. Then, for each row in D we estimate a preliminary attitude score for each topic, by adding the values of the cells of the corresponding seed terms. All users with a score above a certain threshold are labeled with the corresponding topic. In our experiments, we defined a threshold of 0.25, implying that only users who strongly used the seed terms of each topic were labeled.

As result, we obtain $D = U \times T$, where the rank of U and T is two. In our context, each topic is an attitude, the matrix U contains the user-attitude associations, and the matrix T contains the term-attitude associations (transposed). We interpret these associations as probabilities.

2.4.2 Attitude Tendency and Polarity

To characterize attitudes, we calculate two metrics common in the sentiment analysis literature to measure the leaning and amount of sentiment: tendency and polarity [98]. Tendency is defined as:

$$\text{tendency}(u) = P(\text{empathy} \mid u) - P(\text{threat} \mid u), \tag{2.3}$$

where, $P(\text{attitude} \mid u)$ is the association between user u and the corresponding attitude. Note that the definition is analog for terms. For tweets, tendency is defined as:

$$\text{tendency}(\text{tweet}) = \sum_{\text{term} \in \text{tweet}} \text{tendency}(\text{term}). \tag{2.4}$$

Note that tendency values close to zero do not imply a neutral attitude, as there could be non-zero contributions in both topics. To clarify this fact, we consider attitude polarity as the amount of associations to both attitudes, defined for users as:

$$\text{polarity}(u) = P(\text{empathy} \mid u) + P(\text{threat} \mid u). \tag{2.5}$$

The definition for terms is analog. For tweets, polarity is defined as:

$$\text{polarity}(\text{tweet}) = \sum_{\text{term} \in \text{tweet}} \text{polarity}(\text{term}). \tag{2.6}$$

In this way, tendency will allow us to group users/tweets (according to their attitude), while polarity will allow us to measure the intensity of the discussion (how polarized is the attitude).

2.4.3 Lexical Characterization

The previous metrics give an overview of user and tweet attitudes. The next step is to characterize grouped tweets belonging to each attitude according to their tendency. To do so, we use a psycho-linguistic lexicon named "Linguistic Inquiry and Word Count" [119]. LIWC is a lexicon used to study emotional, cognitive and structural components contained in a text. In its Spanish version, it contains 7,515 words classified in one or more of 72 categories. Categories are classified into four dimensions: 1) standard linguistic processes (*e.g.*, articles, prepositions, pronouns, *etc.*); 2) psychological processes (*e.g.*, positive and negative emotions); 3) relativity (*e.g.*, time, verb tense, motion, space); and 4) personal matters (*e.g.*, sex, death, home, occupation, *etc.*). LIWC categories are organized hierarchically, for instance, all words related to the category *anger* are also organized in the categories of *negative emotions* or *affect* words.

We seek to estimate the association of tweets by tendency groups to LIWC categories. After classifying tweets into groups, we estimate how associated the words in LIWC are to each group. Note that specific events may entice a more active discussion by either group, increasing the amount of tweets, thus, we need a way to control the association with these activity patterns. In previous work, this has been done to estimate gross community metrics with z-scores [97, 127]. In our case, the definition is as follows:

$$Z_{lt'} = \frac{P_{lt'} - \mu_l}{\sigma_l}, \tag{2.7}$$

where, $Z_{lt'}$ is the association of LIWC category l with the tendency t', $P_{lt'}$ is the mean of fraction of words in l in each tweet with tendency t', μ_l is the mean of fraction of words in l in all tweets, and σ_l is the standard deviation of the fraction of words in l in all tweets. Hence, this relative metric allows us to compare behavior between groups, by controlling for external variability.

2.4.4 Network Assortativity

The previous definitions capture the behavior in expression, however, the social aspect of Twitter allows to also capture network behavior. We focus on two different networks: the mention network, related to discussion, and the retweet network, related to information diffusion. In both networks, node are users, and links are weighted relations between users. Each node has as attributes its associations to each attitude. In the mention network, a directed link between users u_1 and u_2 exists if u_1 mentions u_2 in one or more tweets. The link weight is the number of times this happens. In the retweet network, a directed link between users u_1 and u_2 exists if u_1 republishes content by u_2. The link weight is the number of times that one user retweets another. These kind of networks are commonly analyzed to understand polarization [34]. To be able to analyze connectivity, we will focus on the Largest Strongly Connected Component of each network.

To analyze the networks structure, we estimate the assortativity coefficient with respect to each attitude. The assortativity coefficient is the Pearson correlation coefficient of numerical attributes between pairs of linked nodes (this numerical attributes are the attitudes given by the model). It measures the similarity of connections in the graph with respect to the given numeric attribute [114]. Hence, the assortativity coefficient measures whether people relations are homophilic with respect to attitude. This behavior is commonly found in networks [8], and it has been documented in Twitter political discussion [34], including in Chile [75].

2.5 Results

Here we present the results of applying the methodology from Section 2.4 to the data set from Section 2.3, covering an entire year of discussion about immigration in Chile.

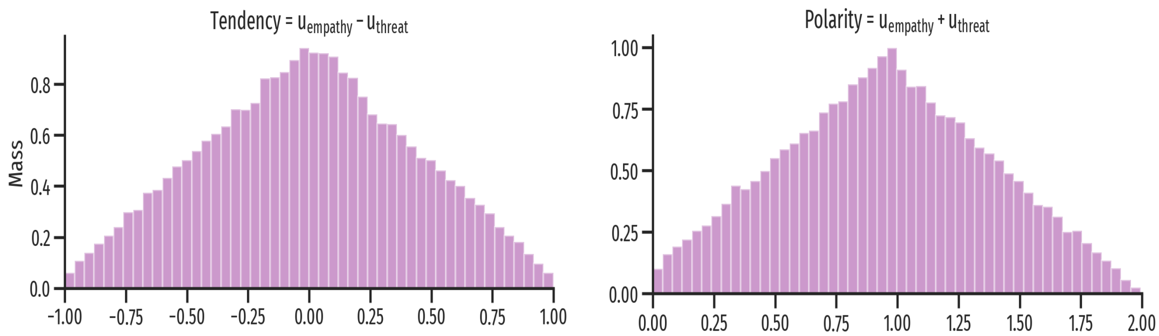
Term Associations Figure 2.3 shows the association of words with each attitude, empathy on the left, threat on the right. One can see that words associated to empathy include “integración” (integration), “salud” (health), and “educación” (education), reflecting their empathetic attitude. Words associated to threat include “delincuentes” (delinquents), “control” (control), and “ilegales” (illegals), reflecting a feeling of threat. Also, empathy group uses the word “Migrantes” (migrants) and threat group uses “Inmigrantes” (immigrants), which can be interpreted as that the empathy group is concerned about the general phenomenon (migration includes emigration and immigration), while the threat group only for the particular phenomenon (immigration).

(a) (b)

Figure 2.3: Most associated words to each attitude according to the TS-NMF model (on the left (a) the empathy group, and on the right (b) the threat group). Note that only single words are displayed, to avoid repetition in n-grams.

Tendency and Polarity Figure 2.4 shows the distribution of tendency and polarity for users. One can see that the distributions are fairly symmetric, with peaks in the center of the distribution.

Figure 2.5 shows the tendency and polarity of tweets during the year under study, estimated using LOWESS. One can see that the tendency trend exhibits two interesting periods, before and after the news about the Leprosy case of an



(a) (b)

Figure 2.4: On the left (a) tendency distribution for users, and on the right (b) polarity distribution for users.

Haitian in July 31th. In the first period, tendency is slightly negative (threat), with an arguably low variability. In the second period, variability increases, and a small negative trend appears, even though at a point in time it reaches its maximum value (*i.e.*, maximum empathy) at the beginning of October. This could be explained by a news event reported in October 6th, about a Colombian citizen that gave birth on the street because a taxi driver expelled her from his car.

It is interesting that both news are related with the Integrated Threat theory and Intergroup Contact theory, respectively. On the one hand, the first event shows the immigrant as a threat, being a possible source of contagion of a disease (Leprosy). On the other hand, the second event shows the immigrant being a victim of violence and discrimination, which arguably makes people more empathetic. Regarding polarity, the trend exhibits a gradual increase in time, with two interesting peaks. The first one reflects the Leprosy case, and the second one reflects the presidential elections, where migration was a common topic in discussion.

LIWC Analysis Figure 2.6 shows the differences of cognitive and emotional categories from LIWC in tweets grouped by tendency: *empathy* contains all tweets

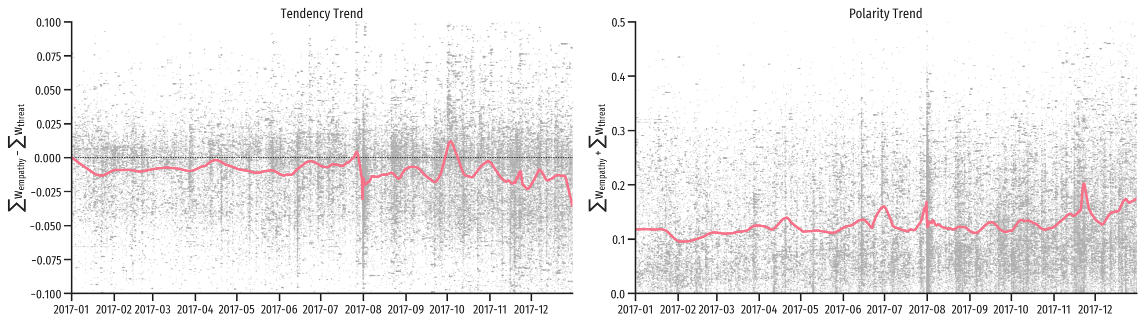


Figure 2.5: Trend distributions (on the left is the tendency, and on the right is the polarity) for all tweets in the data set. Each tweet is a point, the x-position encodes its publication date, the y-position encodes its tendency or polarity. The line is the LOWESS interpolation of tendency and polarity.

with tendency $\geq$ 0; *threat*, otherwise. For each category and group, we estimated the z-score for all tweets each month. As a general observation, one can see that both groups tend to have opposite behaviors. For instance, tweets in the empathy group are positively associated to the *sociability*, *family*, and *positive emotions* category more than tweets in the threat group. Conversely, tweets in the threat group are positively associated with *money*, *job*, and *inhibition* categories. This could be explained by the threat theory, as immigrants can be perceived as an economic threat and labor competition. Also, inhibition category can be interpreted by the desire to prohibit the arrival of more immigrants or to prevent them from accessing social benefits.

Mention and Retweet Networks The largest SCC of the retweet network has 1,239 nodes and 6,441 edges, while the largest SCC of the mention network has 1,868 nodes and 10,201 links. Figure 2.7 visualizes both networks using Hierarchical Edge Bundling [90]. This method allows us to make explicit the adjacency relations between users, as similar edges are bundled to decrease visual clutter. In the Figure 2.7, each link is colored according to tendency of the source node

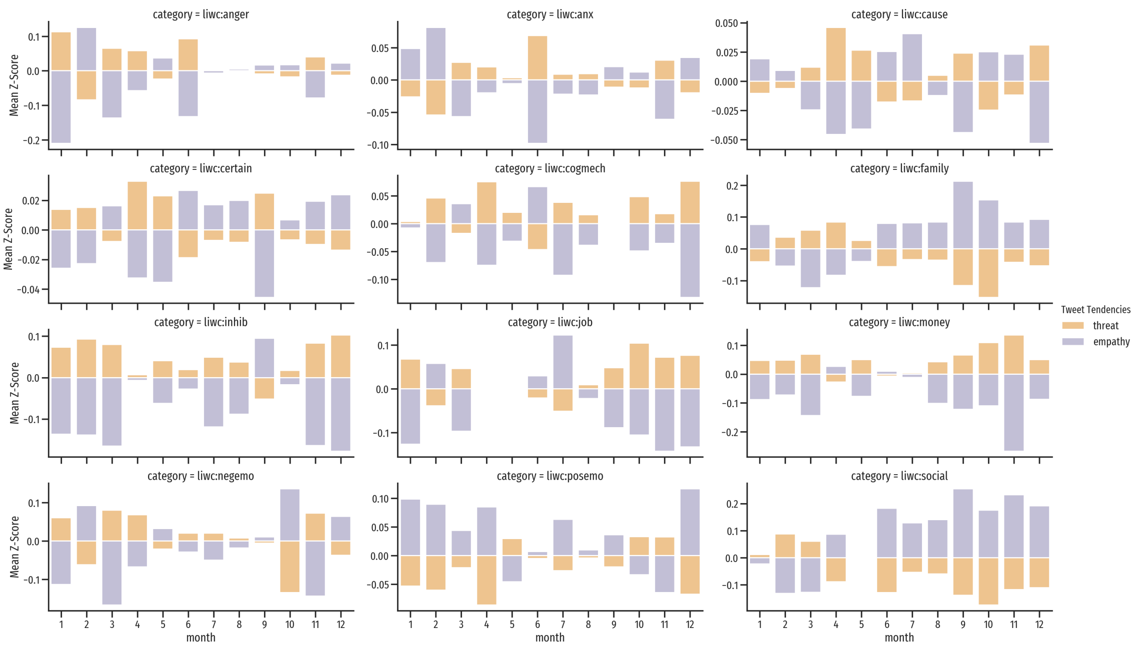


Figure 2.6: Association between attitudes (*empathy* and *threat*) and LIWC categories, per month. Each bar represents the association between groups, estimated with z-scores of fraction of words from each LIWC category and all other words. Purple bars indicate empathy associations, orange bars indicate threat associations.

(purple: empathy group, orange: threat group). Note that the visual encoding makes explicit the community structure in the retweet network and the heterogeneity of the mention network.

The assortativity coefficient for the retweet network are 0.26 (empathy) and 0.14 (threat), implying that homophilic behavior exists, but it is not as strong as in other topics (for instance, the discussion about abortion in Chile is greater [75]), and it is not equal in both groups. As hinted by the visualization, in the mention network the results are small: 0.06 (empathy) and 0.08 (threat). Thus, the retweet network is more segregated than the mention network. This could be explained because retweets are expected to be seen by all followers, and are a key factor in information diffusion, while mentions and replies are not. For instance, one user

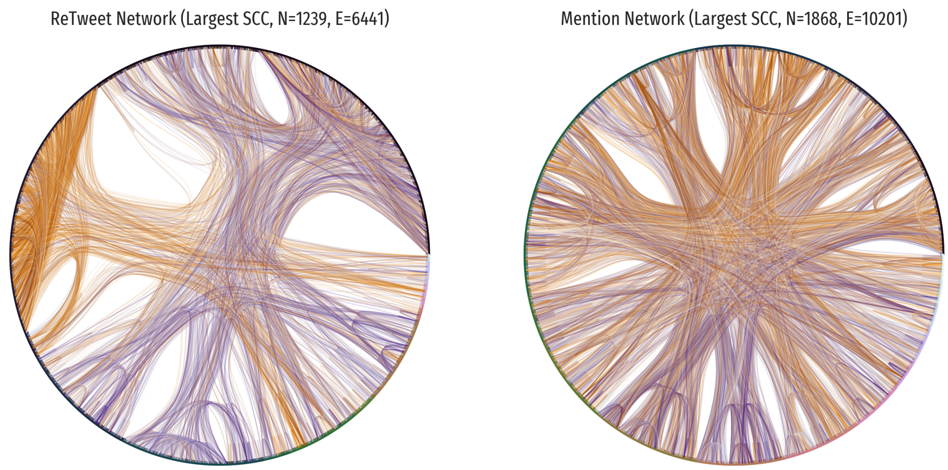


Figure 2.7: ReTweet Network (left) and Mention Network (right). Each node is a circle in the outside, sorted according to the connectivity patters to other nodes. Edges are lines that join nodes, where color is the attitude of the source node (purple: empathy, orange: threat). This encoding allows to group edges that are similar in terms of connectivity between groups.

may send tweets to another holding an opposite position, but if there is no reply, then the interaction is not meaningful.

2.6 Discussion and Future Work

Migration is a controversial issue in Chile, and, although there are some studies about Chileans attitudes toward immigration [101, 27], they do not cover recent migration patterns. To complement knowledge about this topic, we defined a way to classify and measure attitudes, enabling to study the dynamics of perception with respect to immigration and performed a descriptive study of how immigration is perceived in Chile, according to Twitter discussion.

Our results may inform policy and intervention design, as it quantifies how people feel and communicate with respect to immigration. This is relevant, as there exists several contact strategies to improve relationships between social groups [122]. For instance, the discussion we analyzed is mostly targeted at Haitian migration. A majority of them is from Afro-Haitian descent, an ethnicity that was almost non-present in Chile.

There are two key aspects that need further exploration, and that limit the scope of our results: the representativity of Twitter, and the validation of the TS-NMF model. In terms of representativity, Twitter is a biased sample of the population [5]. As such, our results only cover this sample, even though it is not know to which degree nor to which sub-populations it represents. Having these biases into account will surely improve the interpretation of results. However, one aspect that needs to be considered is that Twitter is within the most popular applications in Chile [73], and that it reflects some cultural aspects, such as the country's centralization [74]. In terms of validation, the lack of ground truth or approximate measures of the problem stands in the way of effectively measuring the model accuracy, leaving us only with a qualitative evaluation.

Besides working on the limitations of our approach, there are two lines of future work that we devise. On the one hand, it would be relevant to understand the relationship between attitudes and actual presence of immigrants in a place. This would provide a way to measure real and imagined threat attitudes [96]. On the other hand, there is a potential influence of news events in attitudes. Given the rise of *fake news* and *post-truth* media, this would provide a way to measure the effect of such phenomena on how people feel with respect to a specific issue, migration in this case.

2.7 Conclusions

In this study, we have characterized attitudes toward immigration by locals in Chile. We used a semi-supervised topic modeling technique (TS-NMF [107]) to identify attitudes grounded in two social theories, the Intergroup Contact Theory [3], and the Integrated Threat Theory [141, 112]. Then, we measured differences in attitudes using psycho-linguistic lexicons and interaction networks. As result, we found consistent behaviour with respect to social theory. There is still work to do in the evaluation and representativeness of our model, including the definition of a suitable ground-truth perception to validate our proposal. We believe our results help to inform the design of public policy and interventions to improve relations between groups in a country.

Chapter 3

STUDY 2. ATTITUDES TOWARDS MIGRATION IN A COVID-19 CONTEXT: TESTING A BEHAVIORAL IMMUNE SYSTEM HYPOTHESIS WITH TWITTER DATA

3.1 Introduction

The COVID-19 outbreak implied many changes in the daily life of most of the world's population for a long time, prompting severe restrictions on sociality, as the main policies to avoid contagion consisted of maintaining social distance and establishing long periods of confinement, together with various physical protection and hygiene measures. This evokes a more pronounced reaction to avoid contagion, consisting of all changes triggered by the threat of an infectious disease [110, 135, 136, 137], thrived by the emotion of disgust [146, 147, 103]. These changes could involve changing perceptions of others and the outside world, along with psychological and behavioral responses.

The latter is explained by the activation of the Behavioral Immune System (BIS) [137] motivated by the disgust to pathogens [146]. The BIS is described as a set of mechanisms that help humans and other species to defend themselves against pathogens by detecting them in the immediate environment and facilitating their avoidance before they come into contact with the organism. This response has many implications for humans, such as potentiation of disgust, avoidance of social interactions, establishment of behavioral norms that limit the likelihood of infection, and discriminatory or prejudicial unsociable behavior (including ethnocentrism and xenophobia). Thus, a relevant dimension of BIS is the enhancement of prejudice and discrimination towards those perceived as marginalized groups, including (but not limited to) migrants, defined as people living in a country other than the country of their birth.

From an evolutionary perspective, in ancestral settings, interactions with outsiders posed a high risk of disease transmission, as individuals may not have antibodies against foreign pathogens [111]. Evidence for this remains sparse and mostly uses surveys to measure both xenophobia and pathogen avoidance. Some of this evidence shows that feelings of vulnerability to disease increase negative reactions to foreign groups [50, 111]. However, Hruschka and Henrich (2013) [92] found no evidence to support that hypothesis. Instead, they found that government efficacy is more relevant in explaining in-group favoritism (related to xenophobic tendencies) than disease avoidance. Such discriminatory and prejudiced behavior in times of pandemic is nothing new. Throughout history they have emerged as threat and fear in populations increased, for example; during the Black Death in the 14th century, there was persecution of Jews [104], later in the 1980s, during the HIV outbreak, lesbian, gay, bisexual, and transgender communities suffered great stigmatization and social exclusion [12], and more recently, the Ebola outbreak was labeled as an "African disease" [40]; again stigmatizing this part of the population.

Many studies have focused on finding the factors that influence attitudes toward migration, because of the consequences that they have on social cohesion [91, 95, 96, 140] and on migrants' marginalization, affecting not only individual psychological and socioeconomic factors [24, 138], but the relationship between migrant population and locals, with emerging conflicts that in many cases lead to racism and xenophobia [21, 88, 122]. One theoretical framework that has been developed to understand such attitudes is the Integrated Threat Theory [112, 141]. This proposes that under certain conditions, contact between local and immigrant populations (is generalized to distinct groups) generates perceptions of threat and risk, fostering discriminatory and exclusionary attitudes on the part of the local population. The concept of "contact" is not limited only to physical contact, but also encompasses indirect, imaginary or virtual contact [4, 37, 44, 153]; and the type of perceived threat has been related mainly to competition for jobs, public services, and economic factors in general [24, 138]; however, few studies have focused on the threat produced by pathogens in pandemic contexts.

In this study, we propose to use the information that people post on Twitter about migration as a reflection of their attitudes towards this issue. It is common to find reactions and attitudes through posts in these platforms, where people express their ideas and opinions voluntarily. Studies that have used Twitter reveal socio-cultural characteristics of users or societies, including the influence of culture in personal actions, political polarization [63], personality traits [128], personality differences between democrats and republicans [144], but also the level of integration of immigrants in a city [100] and attitudes in response to triggering events, such as terrorist attacks [39]. In this way, our proposal would complement traditional methods for obtaining information about attitudes toward immigration, which have been measured mainly through qualitative case studies or context-specific surveys for each migration context [85, 116, 125]. Thus, we contribute not only with a dynamic approach to attitudes, but also through an easily accessible and low-cost data source.

Currently, the number of international migrants worldwide has reached 281 million, equivalent to 3.6% of the world population [54]. In Chile, this number amounts to 1.7 million people (7.2% of its population) [52]. Our research seeks to measure attitudes towards migrants before, and during the most critical moments of the pandemic in Chile, using Twitter [59]. In this way we could shed light on how the BIS is operating in a society that has seen its immigrant population grow dramatically in recent years, and has been jointly affected by the global Covid-19 pandemic.

The aim of this study is to test one of the core statements of the Behavioral Immune System: that people tend to enhance their rejection of minorities and foreign groups under the threat of contagious diseases. Specifically, we expect an increase in the number of Twitter users with threatened attitudes (or a decrease in the number of users with empathic attitudes), between pre and post-pandemic contexts. Similarly, the number of threatened tweets should also increase (or the number of empathic tweets decrease). We also expect an increase in the number of retweets from the group of users with threatened attitudes between pre and post-pandemic contexts (or a decrease in the number of retweets from the empathic group of users), which is understood as a greater diffusion and scope of its negativity. Lastly, we expect the language use of the threatened user group to reflect a greater concern for the threat of COVID-19 contagion, i.e., they should use COVID-related words more frequently than empaths, especially in the post-pandemic context.

At this point we will also be working under the prediction of the Integrated Threat Theory, but for a specific case of threat produced by the risk of virus contagion. However, since the BIS can be interpreted as a specific strand of this theoretical model, and our study is comparative for a context with and without Covid-19 pandemic, we will refer to the latter in our predictions.

The chapter is structured as follows. Section 3.2 describes the dataset we used in our study, Section 3.3 describes the methodology applied. Section 3.4 shows the results obtained, and finally, Sections 3.5 and 3.6 presents the discussion, limitations, future work and final conclusions of the study, respectively.

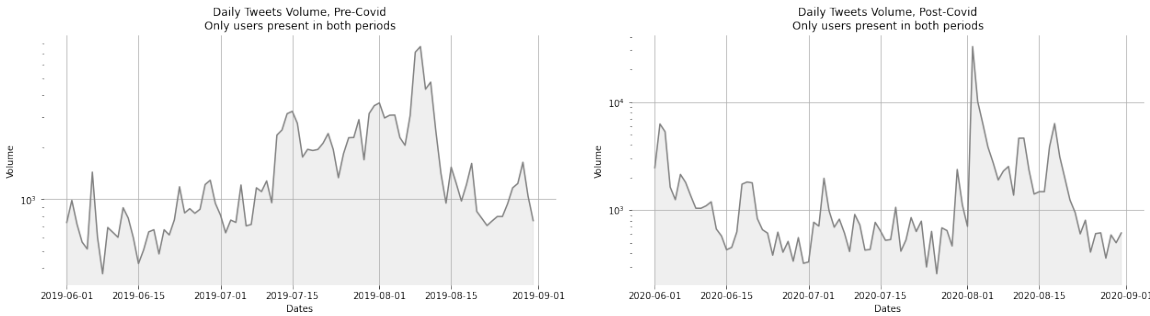
3.2 Dataset Description

In this section we describe the dataset used to measure and analyze attitudes towards immigration.

All of our dataset has been extracted from the Twitter platform. Twitter is a microblogging platform in which each user with an account can create a user name, inform a location, and include a brief personal description, an image, and an URL in their profiles; this information can be fictitious or real according to their own choice. In addition, users have the ability to post Tweets (texts of a maximum of 280 characters), Retweets (share other users' Tweets on their own account), mention others in their own tweets using an identifier (e.g., @username), and quote others' tweets or add comments to them. Moreover, Twitter users can also interact with each other, following other users and viewing their tweets on their own timelines.

To collect the tweets we used a system designed to crawl Chilean tweets [77] which connects to the Twitter Streaming API. Then, to collect those referring to migration we used a list of keywords related to the topic. The query parameters were, for example; "migration", "immigration", "migrant", "immigrant", "foreigner", "borders", etc., and countries of origin with their respective demonyms. Lastly, we pre-process the data to eliminate noise topics such as bird migration, migration in other countries, etc.

Thus, our dataset is composed of 892,487 tweets (192,087 are plain tweets, 700,400 are retweets –RTs–, 109,097 are quotes, and 109,211 are replies) that talk about immigration in Chile, of which; 299,098 correspond to the pre-pandemic period (June - August / 2019), and 593,389 to the post-pandemic period (June - August / 2020). These tweets were written by 213,115 users (82,212 users analyzed in the pre-pandemic period and 130,903 in the post-pandemic period). 21,469 of these users are present in both analyzed periods, and our study focuses precisely on this subset of users since for them we can trace their attitude changes, if any. Thus, the number of tweets analyzed in our study (those that were issued by the group of users present in both periods) corresponds to 312,527 tweets, 146,743 tweets from the pre-pandemic period, and 165,784 from



(a) (b)

Figure 3.1: Daily volume of tweets in the data set (on the left (a) is the pre-pandemic period, and on right (b) is the post-pandemic period).

the post-pandemic period.

3.3 Methodology

In this section we describe how to quantify and classify attitudes towards immigrants. The methodology we applied has two steps, following the works of Freire-Vidal et al. 2021 and Graells-Garrido et al. 2022 [59, 70]: first, we identify the attitude of some users, i.e., we label them as empathic (positive attitude) or threatened (negative attitude). Second, with the subset of users classified, we trained a classifier and extended it to the rest of the dataset using a bootstrap approach [60, 61].

For the **labeling process**, we identified attitudes toward migration based on two social theories. The first one is the "Intergroup Contact Theory”, which proposes that the contact between people from different groups will foster empathy and improve relationships [3]. The second perspective is, the "Integrated Threat Theory”, that states that the contact between people from different groups can cause prejudices and perceptions of threat, which will worsen relationships

[112, 141]. Both theories tell us what to look for when analyzing attitudes toward migrants: either motivated by empathy, in favor of migration; or by threat, against migration. Thus, as in Freire-Vidal et al. (2019, 2021) [58, 59], we will assume that there are two attitudes, which we labeled as *empathy* and *threat*.

Using the fact that an effective mechanism to predict the community to which a user belongs is the use of words that the user shows [22], we generated a list of seed patterns and keywords for each attitude, and from it, we automatically labeled our subset of training users. Thus, we iteratively explored the data set in search of traits that could be assigned to these seeds and keywords. For empathic attitudes we focused on messages expressing that migration is a human right, immigrants are welcome and deserve equal conditions (e.g., "we are all immigrants", "no one is illegal", etc.). On the other hand, for threatened attitudes we focus on messages that express that immigrants are not welcome and/or negatively qualify their arrival or themselves (e.g., "no more immigrants", "illegals take our jobs", etc.).

The terms we choose for the first user labeling lists are not necessarily very frequent, but rather, they are exclusive between attitudes; that is, in such lists, we focus on each chosen term being used primarily in messages that reflect positive or negative attitudes, as appropriate. For the final training and ranking of our study, we use lists that are built iteratively from a small initial list. In other words, we run the classification step with our initial list, and manually check for new terms that are strongly associated with each attitude; if such terms are new and consistent with our first choice of seed patterns and keywords, we add them to the list and repeat the process (the process ends when the associated terms do not contribute new information). Example seeds for each attitude can be seen in Table 3.1.

Table 3.1: Seed patterns and keywords for labeling each attitude toward immigration.

Attitude	Patterns seeds and keywords
Empathy	#bienvenidosachile, #chilesinbarreras, #chilediverso, #noalaxenofobia, #nomasracismo, #pongamonosinmigrantes, #bienvenidosmigrantes, #nadieesilegal, #todossomosmigrantes ...
Threat	#inmigrantesilegales, #nomasinmigrantes, #nomasilegales, #vendepatria, #fuerailegalesdechile, #invasionmigrante, #nomasinmigracionilegal, #inmigracionilegaldesatada, invasión ...

Once we obtain the final lists, we proceed to label the users that match these patterns; in this way, we build our subset of training users for the classifier. In parallel to the labeling we did through the lists of seed patterns and keywords, we manually tag some accounts of institutions (such as the International Office for Migration and the Jesuit Migrant Service), public figures, journalists, and politicians who have publicly expressed their attitude on migration.

This process is done for the data sets of both study periods, considering that during each period the users maintain only one attitude.

For the **classification process**, we use the subset of users labeled as training for the XGBoost classifier that trains decision trees using gradient boosting [30].

The XGBoost algorithm is based on the construction of a model that is a weighted combination of several simpler models. Each simpler model is a decision tree and is called a "weak tree".

Given a training set $D = (x_1, y_1), \ldots, (x_n, y_n)$ with n instances and $y_i \in [0, 1]$ for binary classification, XGBoost searches for a prediction function $F(x)$ that minimises a loss function $L(y, F(x))$.

The XGBoost algorithm uses the boosting technique, which consists of building sequential models in which each model attempts to correct the errors of the previous model [60]. To do so, at each iteration t, a new weak tree $f_t(x)$ is trained to fit the residuals $r_{i,t}$ of the current model $F_{t-1}(x_i)$, i.e., $r_{i,t} = y_i - F_{t-1}(x_i)$.

The goal is to minimise the loss function at each iteration:

$$L^{(t)} = \sum_{i=1}^{n} L(y_i, F_{t-1}(x_i + f_t(x_i)) + \Omega(f_t), \tag{3.1}$$

where $L(y_i, F_{t-1}(x_i) + f_t(x_i))$ is the loss function corresponding to the addition of the new tree f_t to the current model F_{t-1}, and $\Omega(f_t)$ is a regularisation function that penalises the complexity of the tree f_t to avoid overfitting.

To find the optimal weak tree f_t, XGBoost uses a technique known as "split finding", which looks for the best way to split the data into two groups to minimise the loss function. At each node of the tree, all possible splits of the data are considered according to an input variable and a cut-off point, and the split that minimises the loss function is chosen.

Once the optimal weak tree f_t has been found, a learning factor η is adjusted to control the contribution of this tree to the final model. Finally, the weighted weak tree is added to the current model to obtain the new model:

$$F_t(x) = F_{t-1}(x) + \eta f_t(x). \tag{3.2}$$

This process is repeated until a maximum number of iterations is reached or until an acceptable model accuracy is reached.

For our study, the XGBoost classifier receives as an input a feature matrix defined as the concatenation of several matrices, three of them based on the content emitted by the users, and another three based on the interactions between them:

- In the content-based matrix each row represents user i, and each term j can represent a word, hashtag, username, URL or emoji. Thus, a cell (i, j) contains the number of times user i has used the term j in his tweets, in his biographical self-description and in his profile URL.
- For the interactions matrix we consider that homophily can vary or be absent in different layers of interaction [108], so we use the three types of interaction allowed by the platform to build it; retweets, replies and quotes. Thus, the matrix stores in a cell (i, j), the number of times user i has interacted with user j (e.g., if i retweets j once, $c(i, j) = 1$). In addition, for each type of interaction, we consider whether the user has interacted with other users already labeled with an attitude.

Thus, we train the classifier using the set of labeled users to predict the attitudes of the remaining users (defined in the matrix described above). As parameters of the classifier that allow us to avoid overfitting; first, gradient boosting is performed with an early stop, using a validation set of 10% of the training observations. Second, we remove from the feature matrix the seed keywords we used in the training labeling, since our goal is to classify users that do not use these terms in their content.

The classifier gives us for each user u a value $p_a(u)$, which corresponds to the fraction of decision trees that vote for the corresponding attitude a. This value is at $[0, 1]$, and in our case, since we consider two attitudes, we define the score of one attitude as a function of the other $p_{\text{empathy}}(u) = 1 - p_{\text{threat}}(u)$. Then, we apply a threshold ($p > 0.55$) to consider predictions with a number of voters higher than a random choice. We define users who could not be classified into either of the two attitudes as *undisclosed*. For our analyses, we used both p and categorical classification of each user and tweet.

Finally, we manually check profiles that might have been mislabeled by the classifier, focusing on those that are very active/followed in the discussion. We add those manual labels and repeat this step until no obvious inconsistencies are found.

3.4 Results

The following section presents the main results of our analyses, beginning with a general description of the considered users and their attitudes towards migrants before and after the COVID-19 outbreak. Then, we analyzed both the reach and influence of empathic and threatened users by observing their retweets in the same periods. Lastly we performed a vocabulary analysis for both groups of users by attitude, exploring the pandemic-related words that each group mentioned more frequently in both periods.

3.4.1 General description of attitudes of users pre- and post-COVID-19

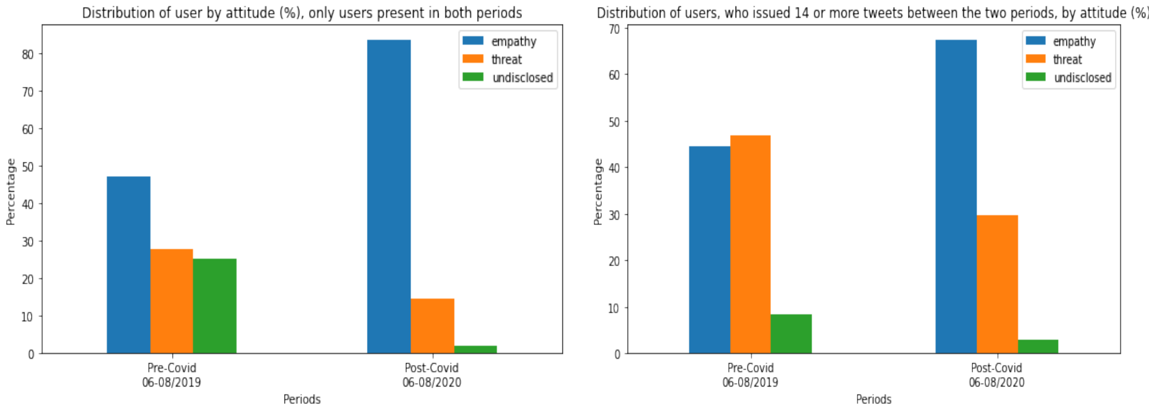
The number of users analyzed in the full period (pre-pandemic: June-August/2019, post-pandemic: June-August/2021) is 231,115, of which 82,212 are users analyzed in the pre-pandemic period, and 130,903 are users analyzed in the post-pandemic period.

For the pre-pandemic period, we have 34,590 users who were classified as empathetic, 16,141 as threatened, and 31,481 as undisclosed. On the other hand, for the post-pandemic period, we have a total of 117,249 users classified as empathetic, 11,678 as threatened, and 1,976 as undisclosed.

In order to study changes in attitude after the Covid-19 pandemic, of the total number of users our analysis only considers 21.469 users whose participation in the Twitter platform was observed in both periods (pre and post-pandemic).

Figure 3.2 (a) shows the distribution of users by attitude. In both periods, empathic are more than threatened users, but pre-COVID we found that less than 50% of the users were classified as empathic, while post-COVID nearly 80% belong to that category. Threatened users, as well as the undisclosed ones, decreased from one period to another. These results are different when we only look at users that have more than 14 tweets about migration and migrants in total, considering both periods (5,312 users). Pre-COVID threatened users were more numerous than empathic ones, but the same change happened when we look at the post-COVID period: the number of empathic users increased, and the number of threatened and undisclosed users decreased, as we can see in the Figure 3.2 (b). We choose the number of 14 tweets to define the group of users who tweet most frequently, because 14 is the average number of tweets users have between the pre- and post-pandemic period.

In addition, as we can see in Figure 3.3, more than half of the users changed their classification from threatened to empathic, and most of the undisclosed users in the pre-COVID period were classified as empathic post-COVID. Since we saw that threatened users decreased post-COVID, we explored the scores the classifier gave to each user, pre- and post-pandemic, aiming to check if they increased between periods. In other words, we wanted to check if threatened and empathic users are more or less threatened or empathic before and after the COVID out-



(a) (b)

Figure 3.2: Percentage distribution of users, classified by attitude for each period of analysis. On the left (a): users present in both periods (n = 21.469), and on the right (b): users have more than 14 tweets between the two periods (n = 5,312)

break. Figure 3.4 shows the average daily score estimated using LOWESS, and we can see empathic post-pandemic users scored more empathic than the same group before the pandemic (from 0.78 to 0.91). Meanwhile, threatened users decreased their average scores from 0.87 to 0.80.

3.4.2 Reach and influence of empathetic and threatened users pre- and post-COVID-19

Our first prediction also stated that the number of threatened tweets would increase or the number of empathic tweets decrease. Tables 3.2 and 3.3 show the numbers of users, tweets, retweets emitted and received, quotes, and replies by group and period. From this, we calculated the ratio of tweets by user, considering the users that were present in both periods. For empathic users, the number of tweets per user was 5.4 pre-COVID and 6.16 post-COVID. Meanwhile, for threatened users, the number of tweets per user was 13.6 pre-COVID and 16.4 post-

Figure 3.3: Number of users classified in each attitude, for the pre- and post-pandemic period, and their attitude changes.

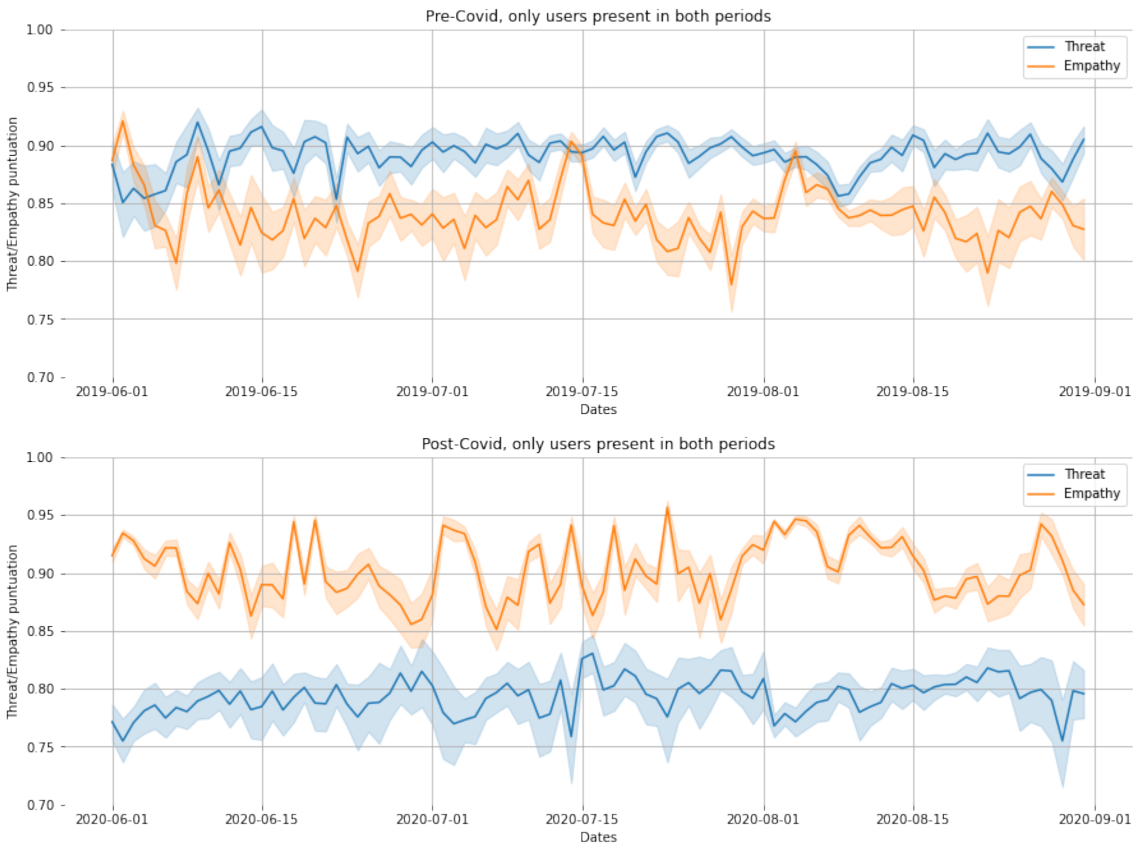


Figure 3.4: Daily distribution of average attitude scores (Empathy and Threat), for both periods (Top: pre-pandemic period. Botton: post-pandemic period).

COVID. Thus, both kinds of users increased the number of tweets they published post-pandemic, but for the threatened users the increase was greater. This result is consistent with the analysis of the complete dataset, where the increase is more prominent for the threatened group (from 9 to 15 compared to the empathic group which went from 3 to 3.4). In contrast, when we analyze only users who have at least 14 tweets in total, although all users increase their tweets, the group of threatened users tweets the most and maintains its average almost constant between both periods (from 27.4 to 28.3), in addition, we observe a notorious increase in tweets per user for the empathic group (from 14.4 to 19.5), and even more accentuated for the undisclosed group (from 5.5 to 20.7).

Table 3.2: Description of the number of users, tweets, retweets emitted and received, quotes, and replies issued by the different groups of users classified according to their attitude, for the pre Covid-19 period.

Period	Pre COVID-19					
Interaction / Attitude	User	Tweet	Retweet emitted	Retweet received	Quote	Reply
Empathy	10,100	54,597	42,047	46,511	6,460	4,602
Threat	5,936	80,883	61,131	83,722	10,447	12,191
Undisclosed	5,433	11,263	7,144	7,630	1,039	2,374
Total	21,469	146,743	110,322	137,863	17,946	19,167

Table 3.3: Description of the number of users, tweets, retweets emitted and received, quotes, and replies issued by the different groups of users classified according to their attitude, for the post Covid-19 period.

Period	Post COVID-19					
Interaction / Attitude	User	Tweet	Retweet emitted	Retweet received	Quote	Reply
Empathy	17,906	110,235	92,967	156,688	12,760	7,515
Threat	3,147	51,439	42,520	51,654	7,460	5,103
Undisclosed	416	4,110	2,877	7,526	505	750
Total	21,469	165,784	138,364	215,868	20,725	13,368

However, this only informs us the users' activity and interactions with others in this environment, not the influence or reach that they have.To check the influence of each kind of user before and after the COVID-19 outbreak, we calculated the ratio of retweets per tweet (dividing the number of retweets, received by each user group, by the number of tweets, issued by each user group) by attitude for both periods. The results for threatened users show a higher ratio before the pandemic (1.04 retweets per tweet), compared to empathic users (0.85). Also, empathics' ratio increases after the pandemics (1.42), remaining over the threatened ones (1.0), the latter remaining stable between both periods. Analysis of the full dataset shows that all users slightly increased their average number of retweets; in contrast to this, for the group of users with at least 14 tweets, only the threatened group shows a decrease in their retweets per tweet between the pre- and post-pandemic period, while the empathetic and undisclosed increase their ratios. This measure considered that there might be users that get more retweets than others because they also tweet more frequently, so it gives us information on the scope and spread of information issued by each user group according to their attitudes.

3.4.3 Vocabulary analysis by attitudes

As a final qualitative analysis, we explored differences in the language use of users with different attitudes. To do so, we calculated the proportional change between the two distributions of words most used by empathetic and threatened users (we took the 50 most frequent words), for the pre- and post-pandemic period.

Figure 3.5 represents the differences in the relative frequency of word use between the two groups of users. In purple color, we have the words used in greater proportion by the empathic group of users, and in orange, those corresponding to the group of threatened users. Here it is possible to observe that in both periods the empathic group uses words such as "children (niños)", "women (mujeres)", "men" (hombres), "persons (personas)" and "people (gente)"; which categorize immigrants in a neutral and equal manner; in contrast with the words "illegal (ilegales)", "tourists (turistas)" and "invasion/invade (invasión/invaden)", used by the threatened users. Furthermore, the group of empathetic users uses words such as "equality (igualdad)", "welcomed (acogido)", "education (educación)" in the pre-pandemic period, denoting their perception of a good reception towards immigrants; in contrast, for the post-pandemic period the words become more accusatory against what does not place them in an equal status (with words such as "inequality (desigualdad)", "discrimination (discriminación)" and "xenophobia (xenofobia)").

Finally, we note that the group of threatened users expresses their threat perceptions around delinquency in the pre-pandemic period, with words such as "arms", "trafficking (tráfico)", "terrorists (terroristas)". Such threat perceptions in the post-pandemic period are replaced by threats of contagion by Covid-19, as only they use words related to the pandemic, such as "infected (infectados)" and "contagious (contagiados)".

To observe more directly the change in pathogen threat perception in both groups of users, we explored the pandemic-related keywords (152) that each group of users mentioned, for both periods. In 3.6 we see that in the post COVID-19 period, empathic users use words related to the disease itself, like "virus", "pandemic (pandemia)", and "COVID-19", while threatened users tend to use words with a more "personal" connotation, like "contagiados" o "infectados" (both mean-

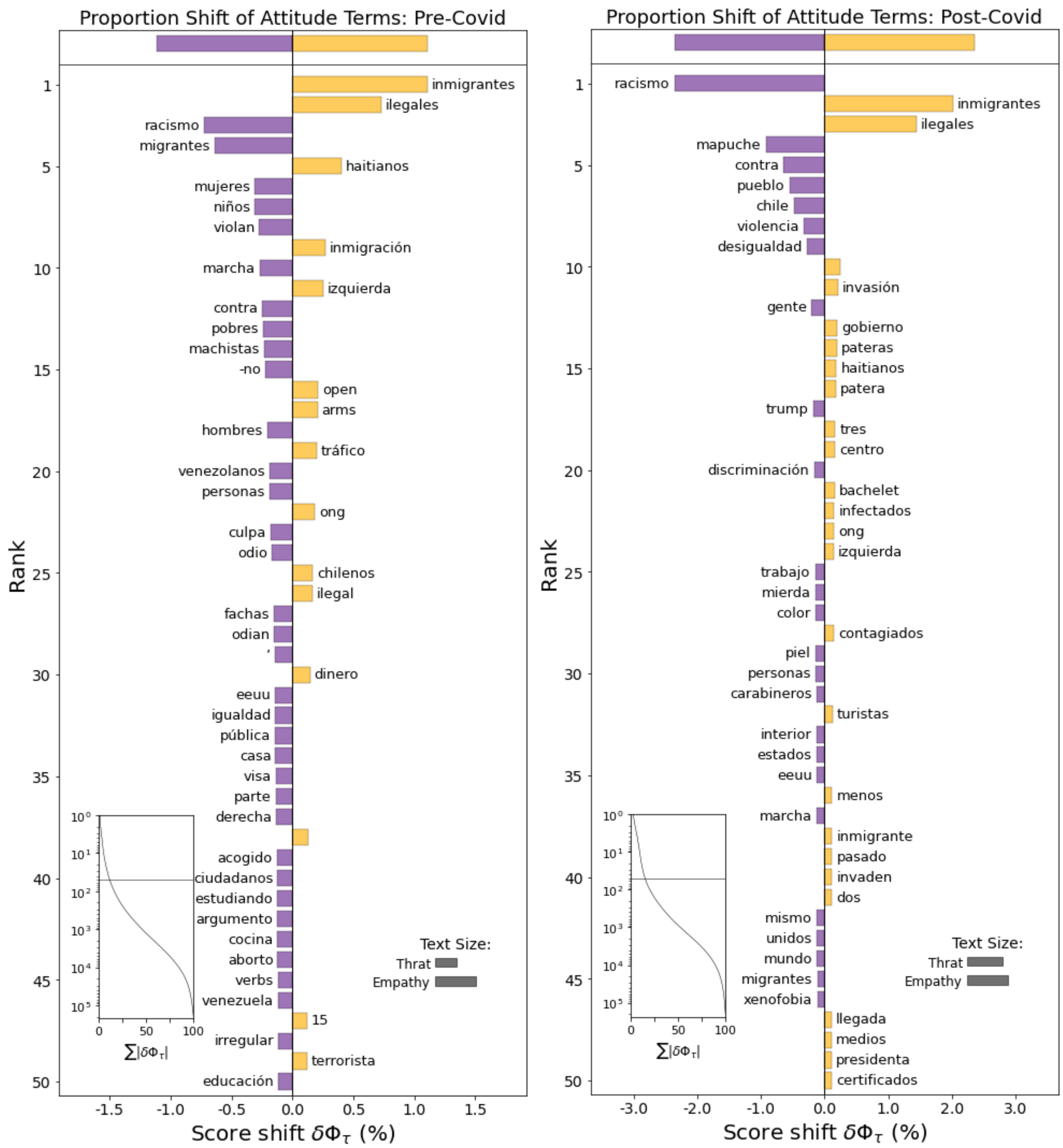


Figure 3.5: Frequency of word usage by each user group (most used words by each group of users according to their attitude, compared to the frequency of use of the opposite group), pre- and post-pandemic (left and right, respectively)

ing "infected").

In addition, the total number of pandemic keywords used by the threatened group is higher than the empathic group for both periods, but for empathics the frequency of use of these words increase from the pre COVID (almost 9 million

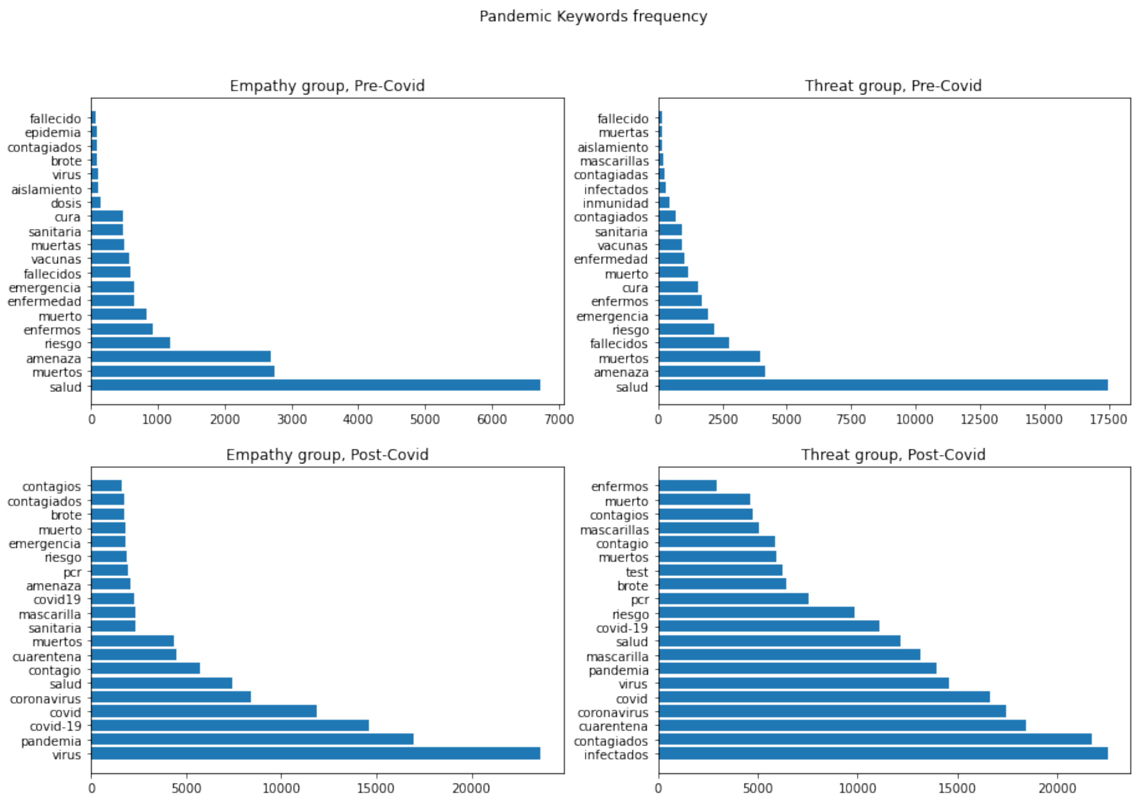


Figure 3.6: Use of pandemic-related keywords in each attitude group, Pre and Post-Covid period.

mentions) period to the post-COVID period (almost 15 million), meanwhile for the threatened group it remained stable. However, when we adjusted for the number of users for each group and period, we saw that the use of these words per user decreases for empathics from 256 to 126 between periods, but it increases for the threatened group (from 1,035 to 1,462 post COVID-19 outbreak).

3.5 Discussion, Limitations and Future Work

One of the core statements of the Behavioral Immune System is that people tend to enhance their rejection of minorities and foreign groups under the threat of contagious diseases [137]. Thus, the aim of this study was to test whether threatened attitudes towards immigrants changed with the COVID-19 outbreak, predicting that these attitudes would increase after the pandemic. Our results partially support this hypothesis.

Specifically, we expected an increase in the number of Twitter users with threatened attitudes or a decrease in the number of users with empathic attitudes, between pre- and post-pandemic contexts. Our results show that, in both periods, empathic are more than threatened users. Threatened users, as well as the undisclosed ones, decreased from one period to another. The undisclosed group stands out for having a very large decrease; this indicates that in the post-pandemic period it was easier to identify user attitudes for our classifier, which can be interpreted as users becoming more clear in their positions and making more noticeable their trends in attitudes. However, these results differ when we only take into account users with more than 14 tweets about migration and migrants in total (average number of tweets users have between the pre- and post-pandemic period), considering both periods (5,312 users): here pre-COVID threatened users are more numerous than empathetic ones. This result is interesting as generally the population with positive attitudes is higher than those with negative attitudes [122, 58, 59]. Given that this group of users shows greater participation in the migration debate and holds more extreme attitudes and positions (which is reflected in the low proportion of users classified as undisclosed), we can see that the proportion of highly participative users with negative attitudes towards the migration debate is greater than those with positive attitudes. Given that in the post-COVID19 period this phenomenon is reversed (the number of empathetic users increases, becoming greater than those threatened) and a similar result is recovered as when we look at all the users analyzed, we could infer that the pandemic dampened the negative attitudes and increased the positive ones for the group of highly participative users.

Similarly, we expected that the number of threatening tweets to increase or the number of empathetic tweets to decrease. The evidence for this is that, for both types of users, the number of tweets per user increased between periods. However, threatened users have a higher number of tweets per user in both periods. Furthermore, although both types of users increased the number of tweets they posted after the pandemic, for threatened users the increase was greater. Thus, we can interpret that the group of users with negative attitudes produces more content than the group with positive attitudes, and their quality of producers is accentuated in the post-pandemic period, which supports the BIS hypothesis. The result is different when we look at the group of highly participative users (who have 14 or more tweets in both periods), since there the average production of threatened users remains almost constant (from 27.4 to 28.3 tweets per user), and the empathic and undisclosed groups show the greatest increase, the latter standing out with an increase from 5.5 to 20.1 tweets per user. This would be consistent with the conclusion that the threatened highly participative group shows a more stable behavior, and it is the empathics who show the greatest change in the face of the pandemic.

Although these results seem to go against the BIS hypothesis, we must consider that a large part of the migrant population worked in essential services (health professionals, supermarkets, pharmacies, garbage collection, deliveries, etc.) during the COVID-19 pandemic; therefore, the local population mostly maintained contact with them (in urban contexts) [49, 67]. Thus, we could interpret that it is this contact that favors the positive perception of migrants, and with it the increase of empathetic users and the decrease of threatened users. With this in mind, it is possible that the Intergroup Contact theory explains this result [3]. The main predictions of this theoretical model are that increased contact between locals and immigrants will promote empathy and understanding between them, and consequently improve and increase positive attitudes towards immigrants.

We also expected an increase in the number of retweets from the group of users with threatened attitudes between pre and post-pandemic contexts or a decrease in the number of retweets from the empathic group of users. To check the influence of each kind of user before and after the COVID-19 outbreak, we calculated the ratio of retweets per tweet by attitude for both periods. Threatened users have a greater ratio before the pandemic, compared to empathic users, meaning that their reach was higher in that period. However, empathics' ratio increases after the pandemics, remaining over the threatened ones, the latter remaining stable between both periods. This measure considered that there might be users that get more retweets than others because they also tweet more frequently, so it gives us information on the scope and spread of information issued by each user group according to their attitudes.

Lastly, we expected that threatened users would use COVID-related words more frequently than empathetic users in the post-pandemic context. We verified this prediction in the first instance, since the group of threatened users is the only one that shows the use of words related to the pandemic, such as "infected" and "infected", in the post-pandemic period. In addition, we found that in the pre-pandemic period this group of users shows threat perceptions of crime (using words such as "arm", "terrorist" and "trafficking"), so that the Covid-19 outbreak may have had an effect on the change of threat perception from crime to pathogen contagion.

Zooming in on the language related to the pandemic, we explored the use of 152 pandemic-related keywords for each group of users. Our results points out that empathic users use words related to the disease itself, like “virus”, “pandemia”, and “COVID-19”, while threatened users tend to use words focused on individuals who pose a threat of contagion (words with a more “personal” connotation), like “contagiados” o “infectados” (both meaning “infected”). Similar to what was found by Y. Freire-Vidal et al. (2019) [58] where the empathetic group of users refer to migration as a general phenomenon, and the threatened group refers to the particular phenomenon that concerns them "immigration". In addition, the total number of pandemic keywords used by the threatened group is higher than the empathic group for both periods. Interestingly, when we adjusted for the number of users for each group and period, we saw that the use of these words per user decreases for empathics from 256 to 126 between periods, but it increases for the threatened group (from 1035 to 1462 post COVID-19 outbreak). This would reflect the fact that for this group of users the perception of threat of infection from the interaction with inmigrant population is higher.

Thus, we have not found clear evidence of increased negative attitudes towards migrants, similar to the findings of F. Rowe et al. (2021) [132]. The fact that we have no validation of the hypotheses derived from the BIS is certainly influenced by a large number of factors internal and external to the population under study. We can mention, for example, that anti-immigrant sentiment was not established in Chile because it is a country that, although it has historically received migrants, its population was very small until 6 or 7 years ago, at which time Chile has experienced a drastic increase in immigration [52]. In addition, the migrant population in Chile is mostly Latin American [52], and this particular population was not directly associated with Covid-19, as was the Asian population, who were the focus of discrimination in different parts of the world [51, 32]. Accordingly, it is possible that the immigrant population was seen by most of the Chilean population as helpers or victims rather than a threat.

Among the main limitations of our work, we recognize the possibility that Twitter only captures the public opinions of a selected segment of the population, whose size and attributes vary according to the region of the country and access to the propper technologies, which ultimately provides a partial portrait of attitudes towards immigrants. Furthermore, given the assumption that attitudes would be constant in each period, by considering the average attitudes of each user's tweets, it is possible that we do not capture the changes in attitudes on a quasi-real scale, and we may be leaving out some fluctuations inherent to the dynamics of the phenomenon. This is especially relevant because there is evidence that points out that attitudes towards migrants are strongly influenced by specific events related to migration [58], such as specific news or the political situation. On the other hand, we do not have other sources of information to analyze other characteristics of the users and their process of attitudes' shaping. This, firstly, because for the purposes of classification, a thematic filter is made to keep the tweets relevant to migration, leaving aside topics that could be important to analyze in conjunction with attitudes towards migration, such as political tendency, socioeconomic level, positions on other associated topics, etc. Secondly, the platform does not provide additional information that identifies users in other relevant dimensions.

Two interesting lines of research emerge from our work; firstly, a temporal analysis of attitudes towards migration in conjunction with a detailed analysis of the most relevant pandemic events in the period studied, and their possible influence on the formation and/or change of these attitudes, and secondly, a geo-referenced analysis of users, in order to cross-reference socio-demographic information (e.g. from the national Census) of users and our measurement of attitudes towards migration.

Despite these limitations, this work presents the novelty of using Twitter to classify and then analyze attitudes towards migration, comparing the pre- and post-pandemic context. This platform allows not only to collect a large volume of data at a relatively low cost, but is also capable of capturing the spontaneous manifestation of different points of view on a given topic.

3.6 Conclusions

In conclusion, the objective of the study was to test whether threat attitudes toward immigrants changed with the COVID-19 outbreak; predicting, according to one of the BIS hypotheses, that these attitudes would increase after the pandemic. The evidence does not go in one direction. On the one hand, in support of the BIS hypothesis, threatened users increased, in greater quantity than empathetic users, their tweet production in the post-pandemic period. While empathetic users increased the reach of their tweets between the two periods, as the ratio of retweets per user increased, remaining over the ratio calculated for threatened users. In addition, Only threatened users show a change in threat perception between the pre- and post-pandemic periods, moving from a concern about crime to one about

Covid-19 infection. And on the other hand, empathetic users use words related to the disease itself, such as "virus", "pandemic" and "COVID-19", while threatened users tend to use words with a more "personal" connotation, such as "infected" or "infected" (both meaning "infected"); making it noticeable that users with negative attitudes towards migration showed perceptions of threat.

On the other hand, threatened users were found to decrease from period to period, with the undisclosed group standing out as having a very large decrease, indicating that users became clearer in their positions and made their attitudinal tendencies more noticeable. However, when considering highly participative users, pre-COVID threatened users were more numerous than empathetic users. The pandemic attenuated the negative attitudes and increased the positive ones for the highly participative user group. It is possible that the Intergroup Contact Theory explains this result, given that most of the services continued to be provided by immigrants (supermarkets, pharmacies, medical and hospital centers, cleaning services, etc.), so that they had to face greater exposure to the virus, thus generating empathy in the local population.

Finally, we can infer that the Covid-19 pandemic may have increased the group of users with positive attitudes on a general level, just as it made the group with negative attitudes more vociferous. This work is a contribution to the studies of the impact of the Covid-19 pandemic in Chile. In addition, it addresses the migration phenomenon and the perceptions that arise from it. Our results allow us to evaluate theoretical models that predict the attitudes of human beings towards the exposed scenarios, as well as to show the feelings and concerns of the Chilean population towards immigration in the pandemic context.

Chapter 4

STUDY 3. A FRAMEWORK TO UNDERSTAND ATTITUDES TOWARDS IMMIGRATION THROUGH TWITTER

Abstract

Understanding public opinion towards immigrants is key to prevent acts of violence, discrimination and abuse. Traditional data sources, such as surveys, provide rich insights into the formation of such attitudes; yet, they are costly and offer limited temporal granularity, providing only a partial understanding of the dynamics of attitudes towards immigrants. Leveraging Twitter data and natural language processing, we propose a framework to measure attitudes towards immigration in online discussions. Grounded in theories of social psychology, the proposed framework enables the classification of users' into profile stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and temporal stance trends. We use a Twitter sample composed of 36 K users and 160 K tweets discussing the topic in 2017, when the immigrant population in the country recorded an increase by a factor of four from 2010. We found that the negative attitude group of users is smaller than the positive group, and that both attitudes have different distributions of the volume of content. Both types of attitudes show fluctuations over time that seem to be influenced by news events related to immigration. Accounts with negative attitudes use arguments of labour competition and stricter regulation of immigration. In contrast, accounts with positive attitudes reflect arguments in support of immigrants' human and civil rights. The framework and its application can inform policy makers about how people feel about immigration, with possible implications for policy communication and the design of interventions to improve negative attitudes.

Keywords: Social Network Analysis; Attitude Classification; Psycholinguistic Analysis; Public Policy; Migration

4.1 Introduction

International migration has emerged as a major divisive global political and societal issue during the 21st Century, with increasing expressions of anti-migration sentiment [29]. Immigration has been portrayed as a major threat to social cohesion, notably during the UK Brexit Referendum and Trump presidential campaign, and drawn attention towards more restrictive migration policies, particularly in Western European countries and the United States [7, 78, 43]. Immigration sentiment is also an essential component for successful migrant integration into receiving societies. Discrimination, intolerance and xenophobia can hinder immigrants' capacity to secure employment, housing and achieve a sense of belonging in local communities, contributing to less cohesive societies [16, 31, 121]. Global initiatives have been established through the United Nations' Sustainable Development Goals (goal 10) [149] and Global Compact for Safe, Orderly and Regular Migration (goals 16 and 17) [53] to tackle anti-immigrant behaviour and thus facilitate migration integration.

The anti-migration sentiment is generally shaped by misconception [29], and social media has become a key channel to spread misinformation, contributing to the formation of misconceptions and manifestation of discriminatory acts against immigrants [132]. However, evidence from experimental study designs have revealed that attitudes can be shifted towards a more supportive view of immigration by explicitly addressing misconceptions [79]. Timely access and understanding of public opinion towards migration is thus critical for tackling misconceptions and understanding shifts in local openness to immigrants [42].

Empirical studies on attitudes towards immigration typically draw on survey questions about existing levels of immigration [132]. While surveys are useful, they are an expensive resource in terms of financial cost, labour and time. Moreover, considerable latency may impact data releases, impairing our ability to regularly monitor changes in migration sentiment, identify and tackle prejudice comments against immigrants. However, we know that anti-migration sentiment and prejudice comments can surge during economic recessions [24] and pandemics [133].

Digital trace data sources can now be used to complement and address some of the shortcomings of traditional survey data. Social media platforms, for example micro-blogging sites such as Twitter, offer a major source of real-time information to understand and quantify attitudes towards immigration. Twitter not only serves as a public forum to exchange opinions and ideas on a broad set of societal issues, including political events [155], abortion legislation [64, 72] and migration [19, 134, 10, 33, 25], it also shapes the opinions of its users [124]. However, Twitter has also enabled the spread of misinformation and negative rhetoric fueling hate speech [152, 20, 80]. Such content has the potential to cause harm to individuals. It often translates into social tension outside the digital world and has played a role in the spread of hate speech during the COVID19 pandemic [133]. With this context in mind, we aim to answer the following research questions: (RQ1) Can we identify, quantify and classify attitudes towards immigration from social network data? (RQ2) What characteristics differentiate the content emitted by users with different attitudes? (RQ3) What emotions and psycholinguistic categories differentiate attitudes?.

Using Twitter data and machine learning techniques, we aim to develop a replicable analytical framework to measure and analyse attitudes towards immigration. Specifically, we propose a framework to: (1) identify users' profile stances of positive and negative attitudes towards immigration; (2) analyse the content and psycholinguistic compositions of these profiles; and, (3) monitor their publication activity rhythm over time. We draw on a sample of 106 K tweets and 36 K users discussing immigration in Chile during 2017 when the immigrant population in the country was recorded to have increased by a factor of seven since 2002 from 105 k to 746 k, with over half of new arrivals occurring between 2012 and 2017 [48].

Our contributions are three-fold: first, we propose a methodological framework to operationalise mainstream theories of social psychology on the formation of attitudes towards immigration using Twitter data. It enables identifying users' positive and negative stances on immigration based on their public opinions and characterising the content and psycholinguistic features of these stances. Second, our proposed methodology reveals how digital traces can be used to complement and augment traditional data sources by enabling understanding of short-term changes in attitudes towards immigration and multidimensional views of immigration sentiment. Finally, our case study provides valuable insights into our limited knowledge of the patterns, experiences and challenges of recently arrived immigrants in Chile. Specifically, our work offers insights into the formation of attitudes towards immigration in Chile during a period of large migration influx, largely related to human displacement in Colombia and an exodus from Venezuela.

The chapter is structured as follows. Section 4.2 discusses the theoretical and conceptual background related to our research work. Section 4.3 describes the dataset used for analysis. Section 4.4 describes the proposed methodology before Section 4.5 presents the results. Section 4.6 discusses the implications and limitations of our work, and Section 4.7 offers some concluding remarks.

4.2 Related Work

In this section, we review theoretical approaches to the formation of attitudes towards immigration. Then, we describe the literature on measuring attitudes based on Twitter data to illustrate the key challenges and advantages of using this data source. Finally, we describe immigration in Chile.

4.2.1 Theories and Measurement of Attitude Formation

Two theoretical models are often used to describe the formation of attitudes towards immigration. These are the Intergroup Contact Theory (ICT) [3] and the Integrated Threat Theory (ITT) [112, 141]. ICT explains how *positive* attitudes form, while ITT describes how *negative* attitudes are created.

The ICT postulates that increased social interaction among people from different population subgroups reduces prejudice and enhances trust. The theory is based on conditions such as common goals for different groups and the benefits of cooperation and support (both formal and informal) to reach those goals. Existing research has contributed evidence to support these arguments [117, 123, 122]. Increased intergroup interaction reduces fear and anxiety that may exist when people interact with individuals from an unfamiliar group [9, 143] by promoting empathy and understanding [14, 142].

In contrast, the ITT predicts that social interaction among people from different groups may lead to perceptions of threat and contempt toward members of the different groups [112, 141]. Two types of threat describe the formation of negative attitudes: symbolic and realistic. In the context of migration, these threats are related to competition in the labour market, to public health concerns from possible diseases, to increased crime and physical well-being, and to perceptions of the size of the foreign group, among others [47, 82], as well as an increased fiscal burden [86].

The common method to measure attitudes toward immigration is to use public opinion survey data. Some examples include the Gallup World Poll, the Pew Global Attitudes Survey, the International Social Survey Programme, the World Values Survey, the Iposos Global Trends, the European Social Survey, and the

Eurobarometer [56]. Data from the Gallup World Poll revealed that in all major regions of the world, people are more likely to want immigration to remain at the current level or increase, rather than decrease, with the exception of Europe. However, there is great variability between countries. In Europe, for example, southern Europeans tend to show more negative attitudes towards immigration than northern Europeans, who show more positive attitudes [43].

Although there is now a large repository of public opinion surveys on migration, some important limitations with these traditional methods arise. For example, existing surveys are mainly from European and North American countries, the number of questions collecting opinions on migration is highly variable across countries, their frequency is coarse, and they are often costly to implement. Thus, it has become a necessity to obtain frequent, comprehensive information covering different regions of the world to understand attitudes toward immigration, as stated in the first objective of the Global Compact for Safe, Orderly and Regular Migration: "to collect and use accurate, disaggregated data to formulate evidence-based policies" [53].

In this work, we propose that the discussion in social networks offers a novel and inexpensive source to complement traditional data systems. It can be collected in near real-time, which would allow for richer analyses grounded on the aforementioned theories. In the following, we describe the potential of the social network Twitter in this regard.

4.2.2 Social Media Analysis in the Study of Human Behaviour

To discuss social media analysis from three perspectives: human behaviour, attitude classification and psycholinguistic analysis. We focus on the analysis of the micro-blogging platform Twitter.

Researchers from different disciplines have used data extracted from the Twitter platform to study psychological and socio-cultural characteristics of communities. Some examples include: personality differences between Democrats and Republicans [144], influence of diurnal and seasonal variability on mood [68], happiness associated with Christianity/atheism [131], political polarisation [34], influence of culture on personal actions [63], prediction of attitudes in response to a trigger event [39], measurement of the level of integration of immigrants in different cities of the world [100], among others. When the analysis is about groups of people, Twitter data highlights its potential use in the detection of hate speech [57], i.e., threatening, harassing or seriously offensive language, as well as to characterise hateful users [130]. The detection of hate speech against immigrants is a particular case of the general framework of hate speech detection. Several studies have used Twitter data to study hate speech against immigrants [19, 134, 10, 33, 25]. C. Arcila C. et al. [25] model and characterise anti-immigrant hate speech on Twitter in Spain. They find that hate speech against immigrants includes Islamophobia, rejection of public support towards the immigrant group, and has a greater presence of offensive than violent language.

In this context, we develop a mixed approach grounded on the theories about attitude formation instead of following a hate speech approach, which is limited to hate but not necessarily opposition/approval or feelings of threat/empathy toward migration. Each formation theory defines an attitude, and, in cases where the classifier confidence is low, we define an *undisclosed* stance to account for participation in the debate without disclosing attitude [156]. Particularly, we build upon our previous work to classify users into attitudes as political stances using a tree-based classifier [72]. As stances, attitudes are not always explicit, and thus, they must be predicted. Two types of features are commonly used for prediction. On the one hand, stance can be predicted using network interactions, based on the assumption that like-minded individuals are more likely to interact [65, 66]. On the other hand, lexical analyses have shown to allow predicting stance as vocabularies within stances tend to have strongly associated words [35, 105], and even other non-textual cues such as emojis [72].

A related area to attitude classification is sentiment analysis. In general, such methods extract the polarity of a text in terms of positive, neutral, and negative intents [154, 118]. These measures may be correlated to attitudes, however, empathetic tweets about migration could speak negatively about discrimination, and threat-related tweets could speak positively about the country, making the interpretation of the results difficult. However, we perform one type of emotional analysis when characterising the content published by users in each attitude. We use the Linguistic Inquiry and Word Count (LIWC) psycholinguistic lexicon [120] to distinguish multidimensional aspects of the migratory discussion. These aspects include multiple emotions and other topical dimensions of expression, which may be compared between attitudes. LIWC usage with social media data includes identifying health issues [41, 87], predicting political sentiment [145], and understanding perception about the transportation experience [150]. The Spanish version has been validated experimentally [129], enabling its consideration in our work.

4.2.3 Immigration in Chile

In the Latin American context, Chile presented the greatest increase in the weight of the immigrant population between the last two censuses [52]. Although immigration is not a new phenomenon in Chile, it has increased dramatically in recent years and has become a relevant issue in the national debate (see Figure 4.1). The immigrant population grew from 1.27% of the total population in 2002 to 4.35% according to the 2017 Census, with a more diverse origin for all immigrants than in previous years. As of 2017, 66.7% of all resident immigrants arrived in Chile from 2010, and 61% arrived in 2015–2017 (considering up to 19 April 2017, the date of application of the census).

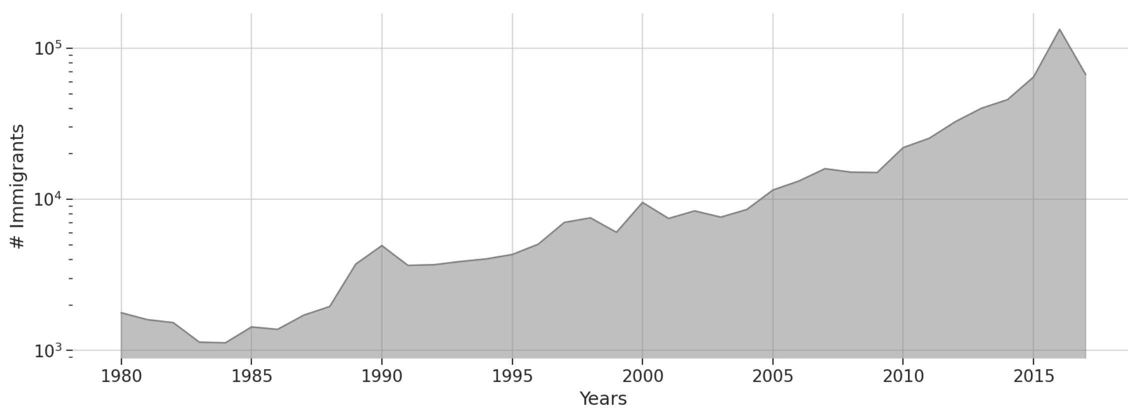


Figure 4.1: Number of immigrants in Chile by year of arrival according to the 2017 Census.

There were two milestones that impacted the national agenda and installed migration as a social concern [38]. First, in August 2016, the Chilean Public Ministry opened an investigation into the possible involvement of an airline in a case of migrant trafficking. Second, in July 2017, the Ministry of Health confirmed a case of leprosy in a Haitian citizen. Leprosy is a disease without records in the country except for overseas islands [11].

A national-level survey in Chile revealed that 57% of nationals agree that the country should take more drastic measures to exclude illegal immigrants, versus 19% who disagree. In addition, 41% of respondents agree that immigrants raise crime rates, compared to 38% who disagree with this statement. In the area of labour competition, 40% of respondents agree that immigrants take jobs away from Chilean-born people, and 36% disagree. This survey was applied in the months of April-May 2017 [28]. Arguably, the questions are more related to the ITT than the ICT. To the extent of our knowledge, there are no other surveys that try to understand how Chileans feel about migration. Hence, this scenario puts Chile as a relevant case study with respect to measuring local attitudes towards immigration. Such analyses may provide complementary knowledge of the attitudes and perceptions of the population, first, by including social theories, and second, by providing fine-grained, dynamic insights.

4.3 Dataset Description

In this section, we describe the Twitter platform and the dataset used to measure attitudes towards immigration.

Twitter is a micro-blogging platform where users report a screen name, a full name (which can be real or fictitious), a location (real, fictitious, or empty), and a small autobiography, among other attributes. Each user publishes posts (called *tweets*) with a maximum of 280 characters. Twitter also allows interaction between its users: users can *follow* others users and to see their tweets in their own timelines. Users can mention other users in their own tweets using a handle (e.g., @username), *quote* other tweets or adding commentary to them, or *retweet* (RT) another tweet to share it with one's audience.

To crawl tweets that discuss immigration in Chile, we connected to the Twitter Streaming API using a system designed to crawl Chilean tweets [77]. The query parameters were keywords related to immigration, e.g., "inmigración" (*immigration*), "inmigrante" (*immigrant*), "fronteras" (*borders*), "racismo" (*racism*), etc.; and origin countries with their respective denomyms. Thus, the dataset after data cleaning is composed by 160,775 tweets (54,252 are plain tweets and 106,532 are retweets—RTs–. In addition, 20,248 tweets are quotes and 19,265 are replies) that are on topic during 2017, written by 36,698 users (see Figure 4.2 for the temporal distribution of content).

The cleaning process ensured that the discussion under analysis was about human migration in Chile. Examples of noise topics were: racism toward indigenous groups in Chile, bird migration, a South American soccer championship, national presidential elections, migration issues in México, the U.S.A., and Spain, among others. Hence, we excluded tweets from users with a reported (or predicted) location different to Chile, as well as tweets in languages other than Spanish.

4.4 Methodology

In this section, we describe how to classify, quantify and characterise the attitudes toward immigration. The methodology is composed of the following steps: theory-informed profile tagging, which enables to pinpoint some users with an attitude; propagation of user attitudes to the rest of the dataset, after training a classifier with the tagged users; and then perform a characterisation of attitudes from the

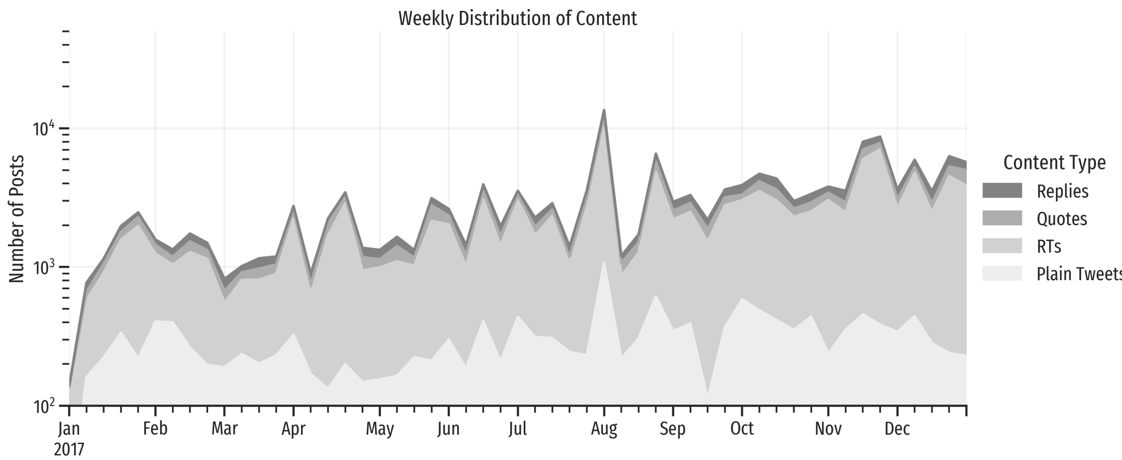


Figure 4.2: Weekly volume of content in the dataset, disaggregated by type of content/interaction.

lens of users, content, psycholinguistics, and dynamics. A schematic diagram of the methodology is shown in Figure 4.3.

4.4.1 Theory-Informed Profile Labelling

Here we define the two attitudes we classify profiles into, and how to label profiles according to these attitudes for classification. In the spectrum of attitudes towards immigration there are two that define two opposite extremes. Usually, these types of categories are named *positive*/*negative*, *in-favour*/*against*, or similar. Here we rely on the social psychology theories described in Section 4.2: the Intergroup Contact Theory (ICT), and the Integrated Threat Theory (ITT). Based on those theories, we name the two attitudes we are classifying as *empathy*, due to the empathy toward immigrants, and *threat*, due to perceptions of threats regarding migration.

Given the potential size of the discussion under analysis, manually labelling the user profiles (or their tweets) into these two categories is expensive and im-

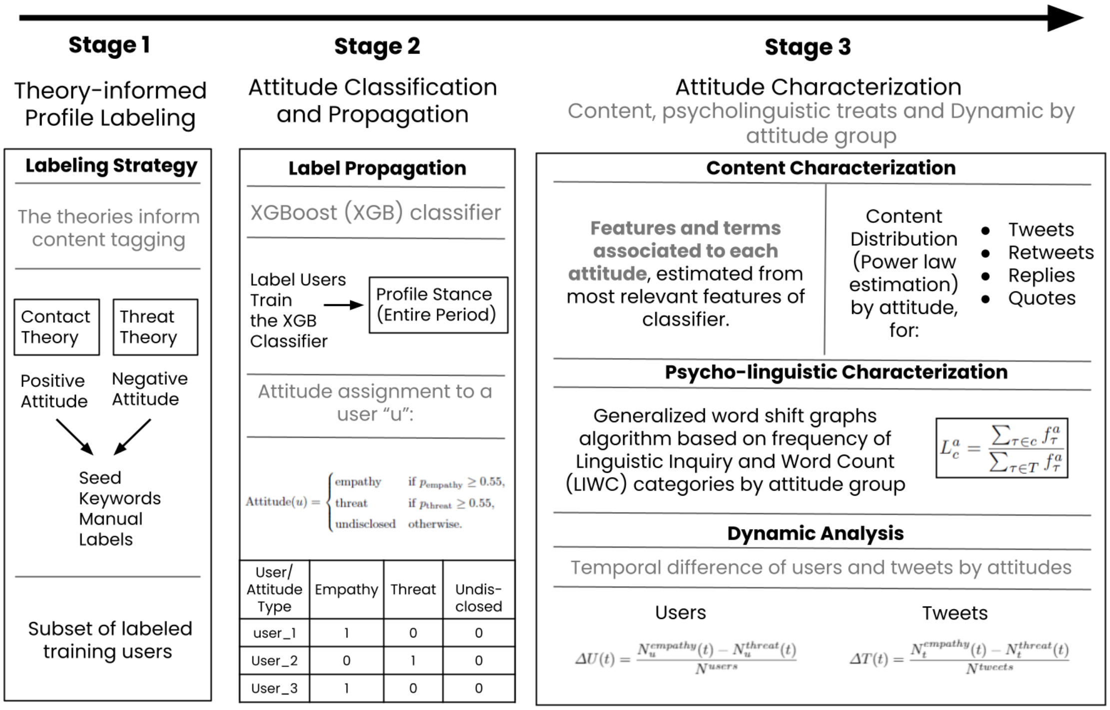


Figure 4.3: Schematic diagram of the proposed methodology.

practical. In view of this difficulty, we predict attitudes using a classifier trained on a labelled subset of the dataset. This subset is labelled automatically from a list of seed patterns and keywords for each attitude, as words are an effective mechanism to predict the community a user belongs to [22, 19, 113]. To identify seeds, we iteratively explore the dataset to seek for features that could be mapped to the *empathy* and *threat* attitudes. In *empathy*, we look for features that indicate that immigrants are welcome and will be received in equal conditions (e.g., "we are all immigrants"). In *threat*, we look for features that show that immigrants are not welcome or qualify them or their arrival negatively (e.g., "illegal immigrants take our jobs"). The labelled terms are not necessarily frequent, instead, they are discriminative, i.e., it is likely that someone in its corresponding category would use the term, and not from the other. The list is built iteratively in the sense of running the first steps from this section up to the classification step, and then exploring the usage of discriminative terms by people in each group to look for other potential seeds. See example seeds for each attitude in Table 4.1.

Table 4.1: Seed patterns for each attitude toward immigration.

Attitude	**Seed Words and Hashtags (and Their Frequencies in Dataset)**
Empathy	#bienvenidosachile (*welcome to Chile*, 26), #chilesinbarreras (*Chile without barriers*, 7), #chilediverso (*Diverse Chile*, 43), #noalaxenofobia (*no to xenophobia*, 126), #nomasxenofobia (*no more xenophobia*, 30), #racismo (*racism*, 1336), #stopracismo (*stop racism*, 48), #nomasracismo(*no more racism*, 10), #noalracismo (*no to racism* 73), #pongamonosinmigrantes (*let us become immigrants*, 67), #todossomoshermanos (*we are all siblings*, 157), #todossomosmigrantes (*we are all migrants*, 1108), #bienvenidosmigrantes (*welcome migrants*, 6), #ningunserhumanoesilegal (*no human being is illegal*, 16), #nadieesilegal (*no one is illegal*, 29), #nohayserhumanoilegal (*no human being is illegal*, 22), #redmigrante (*migrant network*, 105), #interculturalidad (*interculturality*, 218), #díadelmigrante (*migrant's day*, 269), #sinfronteras (*without borders*, 371), inhumano (*inhuman*, 349), multicultural (372), diversidad (*diversity*, 1531)
Threat	#inmigrantesilegales (*illegal immigrants*, 7), #nomasinmigrantes (*no more immigrants*, 41), #vendepatria (*sells homeland*, 5), #estadodecatastrofe (*state of catastrophe*, 44), invasión (*invasion*, 1258), invaden (*they invade*, 62), turba (*group of people generating chaos*, 732), prestamistas (*moneylenders*, 372), narcotráfico (*drug trafficking*, 531), turistas (*tourists*, 2186), fronterizo (*at the border*, 292), enfermedades (*diseases*, 2138), narcotraficantes (*drug dealers*, 291), expulsarlos (*eject them*, 105), echarlos (*to take them out*, 1172), deportarlos (*deport them*, 88)

Next, we label users who match these patterns. Those who are labelled in both attitudes have their labels removed. Note that we assume that in the period under study attitudes do not change. Additionally, we manually label accounts of institutions (such as the International Migration Office and the Jesuit Service for Migrants), opinion leaders, journalists, and politicians that have explicitly expressed their attitude on the issue.

4.4.2 Attitude Classification and Propagation

To predict attitudes, we follow a bootstrapped approach, where we propagate the user labels from the previous step to the rest of the dataset. We use the XGBoost classifier that trains decision trees using gradient boosting [30]. The input feature matrix is the concatenation of several matrices:

- A content-term matrix, where each row represents user i, and each term j can represent a word, hashtag, username, URL or emoji. Thus, a cell (i, j) contains the number of times user i has used the term j in their tweets.
- A profile-term matrix, analogous to the previous one, but this time for the terms contained in the full name and biographical self-description of each user.
- A profile-domain matrix, mapping to each user's home page its main domain (e.g., twitter.com) and their main top level domain (e.g., .com).
- Since homophily may vary or be absent in different interaction layers [108], we consider the three types of interaction separately. Thus, we build three adjacency matrices based on the interactions in the discussion: retweets, replies, and quotes. Each matrix stores in a cell (i, j), the number of times user i has interacted with user j (for instance, if i retweets j one time, $c(i, j) = 1$).
- A user–attitude interaction matrix for each type of interaction, where each cell contains the number of times the corresponding user has interacted with

other users that were labelled with an attitude.

Then, we train the classifier using the set of labelled users. To avoid overfitting, we take two measures. First, the gradient boosting is performed with early stopping, using a validation set of 15% of the training observations. Second, we removed columns from the feature matrix that were used for labelling. This includes the seed keywords for each attitude, as they perfectly separate users from both groups and our goal is to classify users who do not use these terms in their content.

After having trained the classifier, we predict the attitude of the rest of the dataset. For a given profile u, the classifier outputs a value $p_a(u)$ for each attitude a that lies in $[0, 1]$, corresponding to the fraction of decision trees that vote for the corresponding attitude. Note that the value of $p_a(u)$ is not a real probability. Thus, we apply a small threshold to consider predictions with a number of voters higher than a random choice. Those users who cannot be classified are marked as *undisclosed*. As a result of this stage, we assign an attitude to a user u according to the following function:

$$\text{Attitude}(u) = \begin{cases} \text{empathy} & \text{if } p_{\text{empathy}} \geq 0.55, \\ \text{threat} & \text{if } p_{\text{threat}} \geq 0.55, \\ \text{undisclosed} & \text{otherwise.} \end{cases}$$

After predicting attitudes, we manually check for profiles that are highly active/followed in the discussion and could have been mislabelled by the classifier. We add those manual labels and then repeat this stage.

4.4.3 Attitude Characterisation

In the last stage, we characterise each attitude from multiple perspectives. We describe what characterises each attitude from the lens of content (what is published in each category and how?), the lens of psycholinguistics (which semantic categories discriminate the expression of emotions in each category?), and the lens of dynamics (when does each attitude express their opinions?).

Content Characterisation To measure what is published in each category and how, we focus on two aspects of content: the profile features that are most associated with each attitude, and the distribution of tweet vocabulary per attitude. On the one hand, the association of the features deemed important for classification provides insight into which ones are associated with each attitude [72]. However, the classifier estimates plain feature relevance, without any association to each attitude. Hence, for all relevant features, we estimate their association to each attitude using the log-odds ratio with uninformative Dirichlet prior [109], a measure that weights features in a similar way to TF-IDF, with the addition of controlling the variability of frequency. We apply this weighting to three documents, one per attitude (including *undisclosed*). Each document is the column-oriented sum of the feature matrix for all users predicted in the corresponding attitude. Here, we expect to find different terms that express the same concept but are associated with opposite attitudes [113], as well as differences in features from the self-reported biographies, and even in emoji usage [72].

On the other hand, we know that the content generation in Twitter tends to follow a powerlaw [139, 99], that is, their distributions can be described approximately as $P(x) \sim x^{-\alpha}$. Hence, for each attitude, we characterise the volume of content (the number of plain tweets, RTs, quotes, and replies) according to the exponent α of their fitted powerlaw distributions. These exponents enable us to compare if attitudes behave differently in their discussion and interaction mechanisms.

Psycholinguistic Characterisation The previous lens provided differences in the content published by each attitude. However, a direct content-based approach does not discriminate the expression of emotions in each category. To gain a deeper understanding of attitudes in this direction, we use a well-known psycholinguistic lexicon named "Linguistic Inquiry and Word Count" (LIWC) [119]. LIWC was designed to capture emotional, cognitive, and structural components present in text. It is available in several languages. Its Spanish version contains 7,515 words classified in one or more of 72 categories that belong in four dimensions:

1. Standard linguistic processes: articles, prepositions, pronouns, etc.
2. Psychological and affective processes: positive and negative emotions, with subcategories such as anger and anxiety.
3. Relativity: time, verb tense, motion, space.
4. Personal matters: sex, death, home, occupation, etc.

We will focus our analysis on LIWC categories that are possibly associated with some of the factors shaping empathetic and threatening attitudes towards immigration. With regard to *empathy* attitudes, we consider the following LIWC categories as relevant: affective processes, positive feeling and emotions, optimism and energy, humans, social, family, and inclusion. With regard to *threat* attitudes, we consider the following LIWC categories as relevant: anger, anxiety, negative emotions, inhibition, death, body, job, and money. We estimate the association between attitudes and LIWC categories. First, we define the association of LIWC category c to attitude a, L_c^a, as the relative frequency of words in c with respect to all terms in the discussion per attitude:

$$L_c^a = \frac{\sum_{\tau \in c} f_\tau^a}{\sum_{\tau \in T} f_\tau^a},$$

where T denotes the vocabulary, and f_τ^a the total frequency of vocabulary term τ by accounts with attitude a. To explore these associations, we visualise them using Generalised Word Shift Graphs [62]. These visualisations summarise each attitude according to the differences in association between attitudes and LIWC categories, with the aim of describing the emotional and semantic aspects of attitude formation.

Dynamics and Events Our final lens of characterisation aims to understand when each attitude expresses its opinions. Although we assume that attitudes are constant in the time under study, there still could be differences regarding how they are expressed in time in terms of volume. Particularly, we compute the daily difference between the number of tweets and the number of users per attitude. Positive values of this difference indicate a tendency towards *empathy*, whereas negative values indicate a tendency towards *threat*. Then, we compare these time-series with a null-model where we randomly permute the predicted attitudes for users 1 K times. If for a given day, the time-series does not intersect the 95% confidence interval of the null series, we consider it different with statistical significance. Finally, by identifying dates with salient and significant differences, we can explore what triggers the expression of each attitude. In this work, we manually do this through visual exploration.

The proposed methodology provides a full pipeline to characterise users attitudes towards immigration in a micro-blogging platform. Next, we apply this methodology to a set of tweets in Chile during 2017.

4.5 Results

This section first presents the results associated with the classification and characterisation of user attitudes towards immigration according to *empathetic* attitudes and *threatening* attitudes. Next, we analyse the content composition of these user profiles before examining differences in their lexical expressions and temporal fluctuations in attitude sentiment.

4.5.1 Attitude Identification and Classification

To answer our question RQ1, we performed an exhaustive search in the dataset for hashtags and words that could be strongly associated with each attitude (see Table 4.1 for more examples). For instance, the hashtags #bienvenidosachile (welcome to Chile) and #inmigrantesilegales (illegal immigrants) can be directly mapped to each attitude. We found more hashtags in the *empathy* group than in the *threat* group. When tagging threat users, we included words that are not normally associated with threats or negative comments against immigrants. For instance, “turistas” (*tourists*) is not a word commonly associated with a threat attitude. However, in our dataset, tourists tended to be used to express anger about people entering the country on a tourist visa and overstaying.

In total, 3.1K accounts were tagged in the empathy group, and 1.2K accounts were tagged in the threat group (see Table 4.2). The attitude classifier presented good performance, with high precision for both attitudes (0.95 in *empathy*, 0.81 in *threat*) and high recall (0.88 in both). Such good performance is expected, due to a tagging strategy based on a perfect separation of accounts from each attitude. However, this strategy works well identifying a tendency towards attitudes because the keywords used for tagging are excluded from the learning process [72]. Figure 4.4 presents the top-50 features according to their importance

Table 4.2: Performance (precision and recall) of the user profile classifier based on 10-fold cross-validation.

Attitude	Precision (mean)	Precision (std.)	Recall (mean)	Recall (std.)	Labelled Accounts
Empathy	0.95	0.04	0.88	0.13	3118
Threat	0.81	0.16	0.88	0.12	1233

in classification. These features include: retweets to accounts that were be automatically tagged with each attitude, mentions to then-president Michelle Bachelet (@mbachelet), attitude-relevant terms such as cesantía (*unemployment*), delincuentes (*criminals*), inmigración (*immigration*), migrantes (*migrants*), and haitianos (*Haitians*); and mentions to migrant-support institutions as @sjmchile (Jesuit Center for Migration). The remaining features in the top-50 include interactions with authors of viral tweets, and terms relating to specific content, such as a quote by Umberto Eco.

We then used the classifier to propagate attitudes to the rest of the dataset (88% of the total dataset). Table 4.3 reports the resulting classification. Considering the labelling and propagation processes, a total of 71.98% user profiles were classified in *empathy*, 3.24% in *undisclosed*, and 24.78% in *threat*. This is an un-

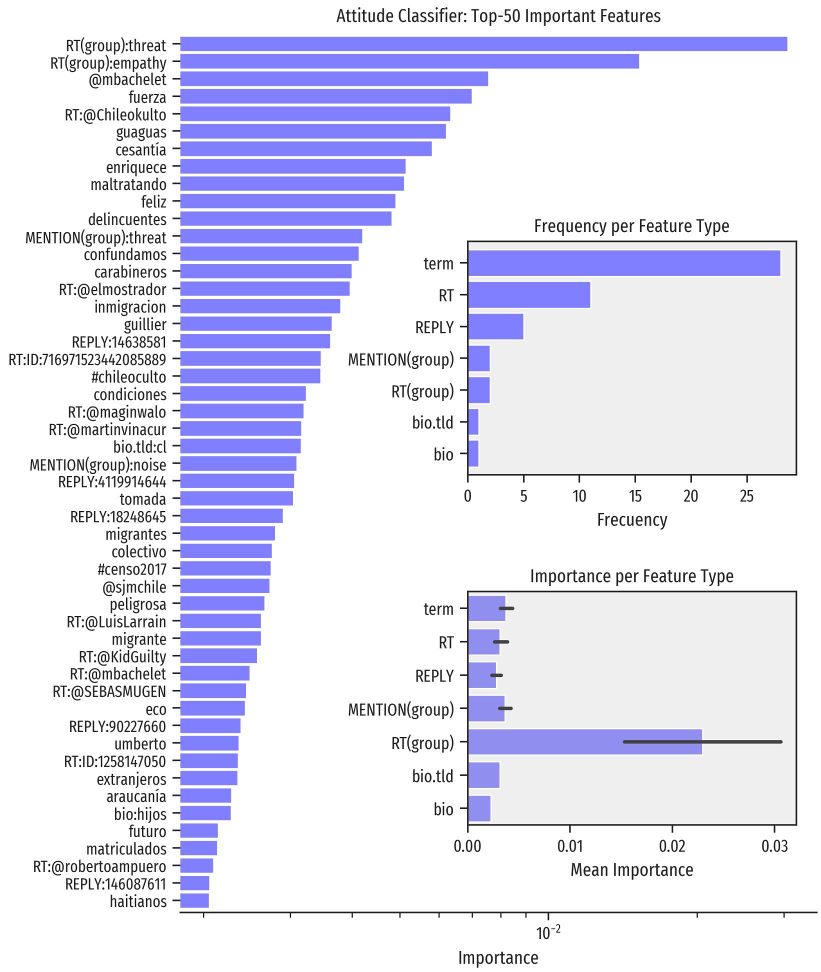


Figure 4.4: Attitude classification: Top 50 most important features.

expected outcome. We expected the distribution of user profiles to be negatively biased given the social divisive nature of immigration in Chile. However, this may imply that negative users are more vocal than positive users and hence there is a perception for negative comments to dominate the public discussion around immigration issues in Chile. Table 4.3 reveals this is the case with the threat group generating considerably more content per user (with a ratio of 3.5 tweets per user, for the empathy attitude group, versus 6.7 tweets per user, for the threat attitude group) than the empathy group, particularly quotes and replies.

Table 4.3: Distribution of user accounts and tweets per attitude and account type.

Attitude	**Accounts**	**Total Tweets**	**RTs**	**Quotes**	**Replies**
Empathy	% 71.98 (26,414)	% 57.46 (92,374)	% 61.94 (65,989)	% 45.62 (9238)	% 32.81 (6320)
Undisclosed	% 3.24 (1190)	% 4.52 (7261)	% 1.44 (1531)	% 1.87 (379)	% 3.61 (695)
Threat	% 24.78 (9094)	% 38.02 (61,140)	% 36.62 (39,012)	% 52.50 (10,631)	% 63.59 (12,250)

4.5.2 Content Characterisation

To better understand the discussion driving the expression of positive and negative attitudes towards immigration and answer our question RQ2, we performed content analysis producing a characterisation of the most frequently used terms associated with each user stance profile. We weighted each term using log-odds ratios with an Uninformative Dirichlet Prior [109]. Figure 4.5 displays the top-30 attributes per attitude, revealing that own-attitude retweeting behaviour is the most prominent feature for *empathy* and *threat*. This is arguably expected due to homophilic interactions [108], a phenomenon that has been observed in political discussion in Chile [76]. However, care must be taken when interpreting the result, as this reflects interactions with accounts that were pre-labelled only. A plausible explanation is that highly popular accounts that can be labelled through methods tend to make their positions explicit.

Neutral terms, such as migrante (*migrant*) and migración (*migration*) and haitiano (*Haitian* as a singular noun) feature prominently in the *empathy* group. Emojis are typically used to express political attitudes [72] and we observe that the red heart emoji is associated with *empathy*, possibly as an expression of solidarity or liking of immigrants-related content. Content related to news items also feature prominently in empathetic Twitter content. Ranking in the top-50 items, we identified the case of Richard Joseph: a Haitian who saved a person that fell from a nine-story

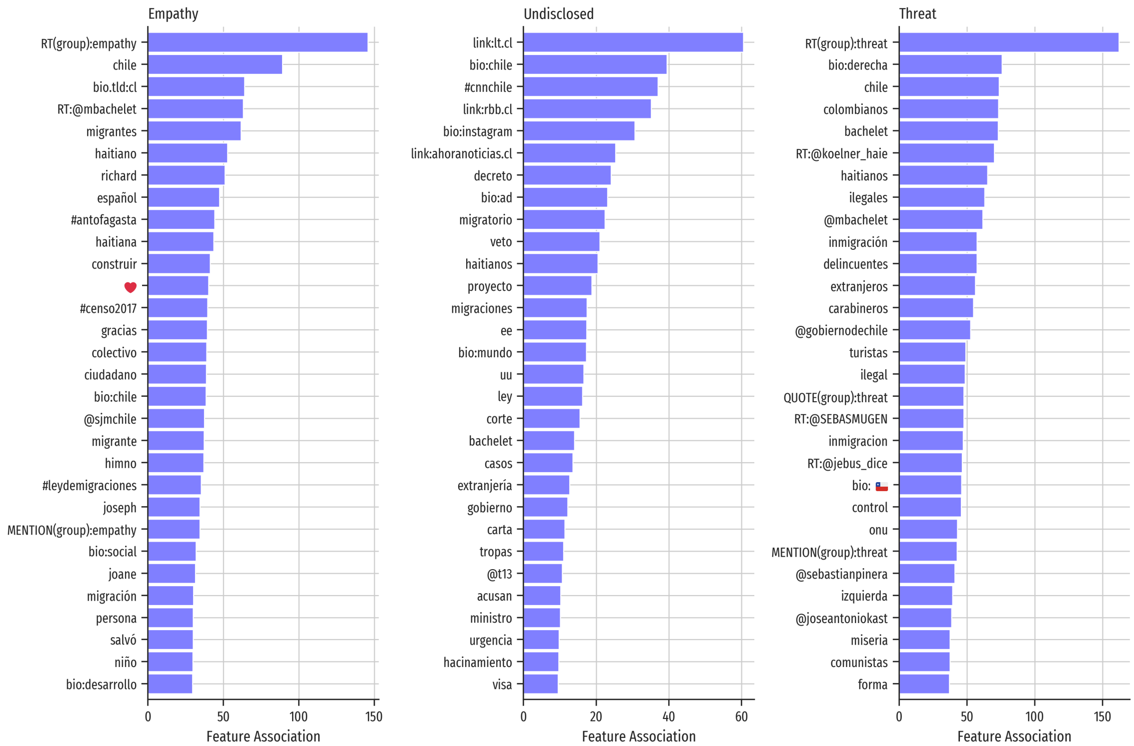


Figure 4.5: Top-term associations to each attitude (or lack thereof), estimated using log-odds ratio with uninformative Dirichlet prior.

building; and, Joane Florvil: a Haitian mother who died after being arrested by the Chilean police under unclear circumstances. We also identified political-based debates, including retweets of Michelle Bachelet's and mentions of a debate of a national migration law (#leydemigraciones).

In contrast, words such as inmigración (*immigration*), extranjeros (*foreigners*) and haitianos (plural, *Haitians*) are often used by the threat group. Immigration placed the focus on migration as a process and concerns about its potential implications for the national health, education and labour market systems, rather than on understanding the individuals themselves. The use of foreigners and Haitians may be used to draw a clear distinction between the "we" and "us". The use of the Chilean flag also features prominently among the *threat* group, arguably associated with expressions of nationalism. *Threat* users also tend to identify their alignment with right-wing views, including “derecha” (right-wing) in their biographies; tagging right-wing politicians, such as Sebastián Piñera (@sebastianpinera, current president of Chile) and José Antonio Kast (@joseantoniokast, extreme right-wing presidential candidate); and, posts about left-wing parties (“ìzquierda”) and the Communist Party (“comunistas”). Overall, the results highlight a strong association between anti-migration views and conservative political ideologies. This is despite not including any explicit political keywords in the seed list for our stance classifier.

We did not observe incidence of the seed words (used to train the classifier) in the content analysis and relevant features associated with each attitude; since only one of them appears within the top ranking (tourists, in the threat group). Note that the total number of training seed words is small (56 in total; 34 for the empathy attitude, and 22 for threat) and their frequency of use is mostly small (see Table 4.1 for some examples).

To identify and quantify differences in the diffusion of content generation, by both groups of attitudes, we analysed the respective distributions of the number of tweets, retweets, quotes and replies. Because other studies suggest that these measures of content generation follow a powerlaw distribution [99, 106], we fit powerlaw regressions on the distributions of the aforementioned measures. Figure 4.6 reports the results. Theoretically, a symmetrically diagonal line would indicate perfectly equal spread generation of content. A perfectly horizontal line would indicate high concentration of the content generation by a single user. Figure 4.6 reveals a relatively symmetrical distribution in the generation of content for empathetic users across tweets, retweets, quotes and replies. This contrasts with the distributions associated with the threat group that displays higher concentration of content generation in a small number of users, with consistently lower powerlaw exponents to empathetic users. This is particularly prominent for retweets suggesting a high degree of interaction within the social network of threat users. These findings confirm our previous interpretations that while our dataset includes a small number of threat users, they are more vocal and a minority of these users tend to generate a comparatively larger amount of content than empathy users.

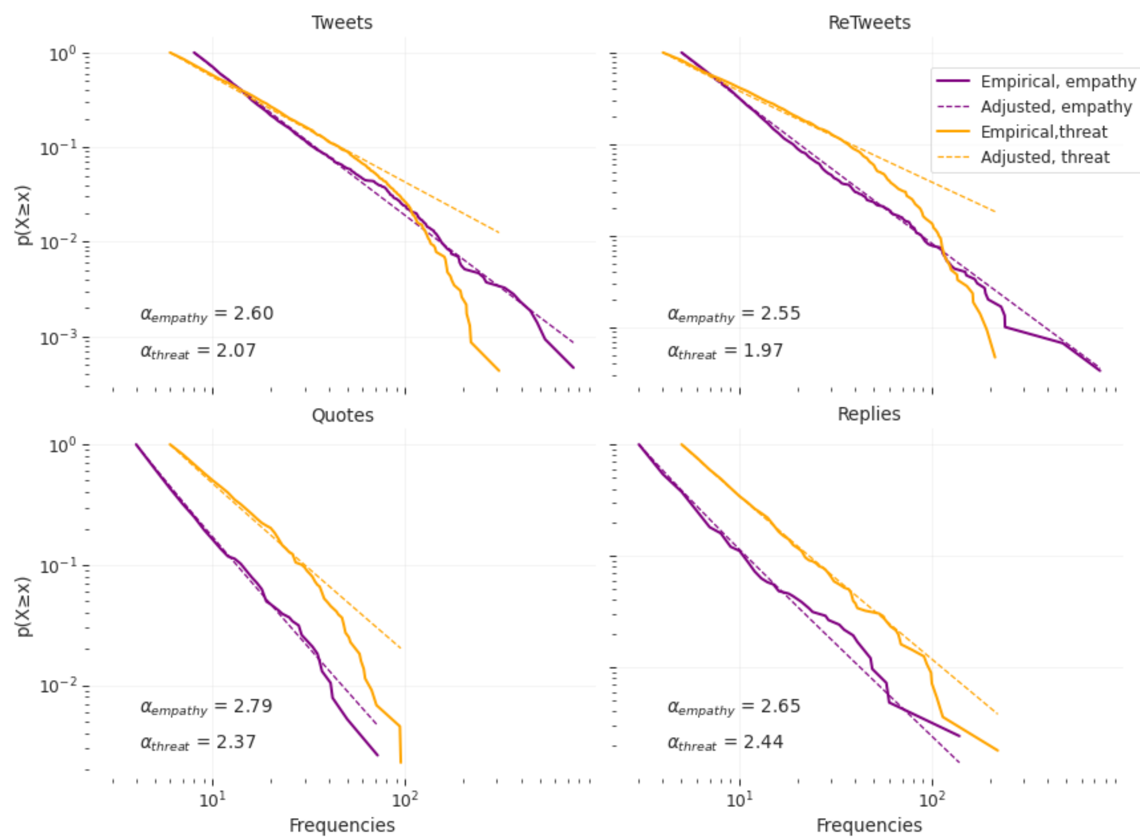


Figure 4.6: Powerlaw distributions for content metrics.

4.5.3 Psycholinguistic Characterisation

We also analysed the manifestation of cognitive and emotional structures in the text based on the LIWC lexicon and answered our question *RQ3*. Figure 4.7 displays the differences in the relative frequency of the LIWC categories between each group, i.e., *empathy* and *threat*.

The results reveal that categories linked to social, posemo (positive emotions), school, comm (communication), humans, optim (optimism and energy), family, affect (affective processes), incl (inclusion), and posfeel (positive feelings) are found more often in the *empathy* group. These are all psycholinguistic concepts used to describe empathetic attitudes associated with contact theory hypotheses. Conversely, categories related to motion (defined by words such as move, walk, go out), othref (reference to other people), present, negemo (negative emotions), death, inhib (inhibition), anger, money, anx (anxiety), and job are more commonly found in expressions used by the *threat* group. This is consistent with the threat theory, as immigrants can be perceived as an economic threat and labour competition.

4.5.4 Dynamics and Events

Finally, we examined changes in user engagement with the migration debate over time. We estimated the daily difference in participation rate between user profile stance; that is, proportion of users in each stance group actively engaging with Twitter content about migration (see Figure 4.8). A positive difference indicates greater engagement among *empathy* users than *threat* users, with negative values denoting greater engagement among the latter. The resulting differences are small, though notable discrepancies exist on specific dates, particularly in

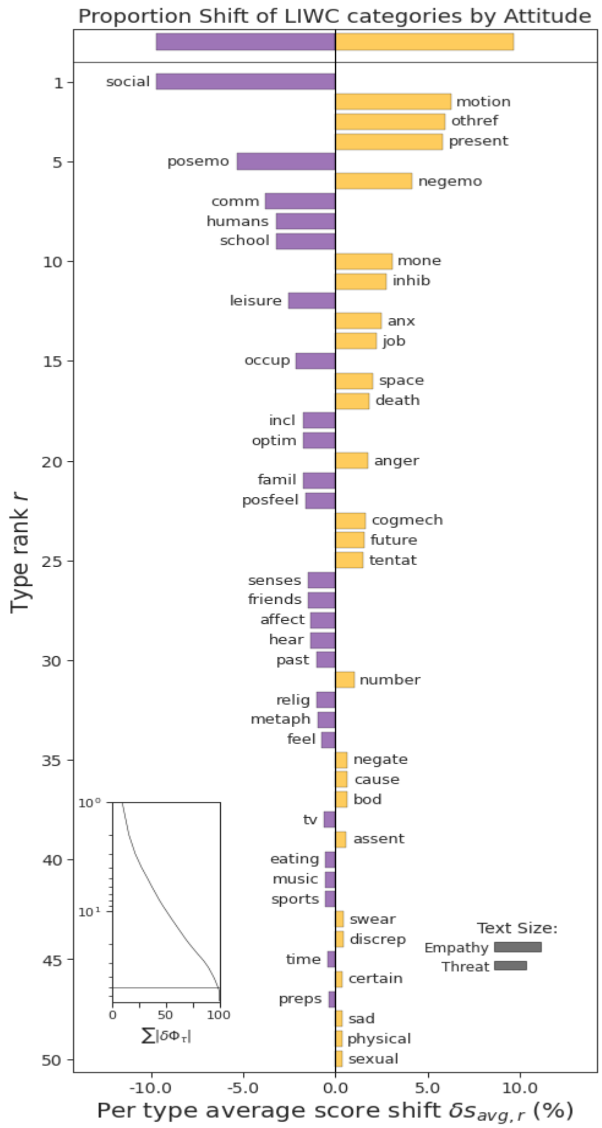


Figure 4.7: Proportion shift of LIWC categories in tweets grouped by *empathy* (on the left in purple) and *threat* (on the right in orange) attitude.

the second half of 2017. Peak discrepancies in engagement coincide with key migration-related events that received mass news media coverage:

- 19/04: The Chilean Census was carried out. Tweets regarding foreign interviewers went viral (positive trend spike).
- 01/07: The National Institute of Human Rights filed a complaint against a

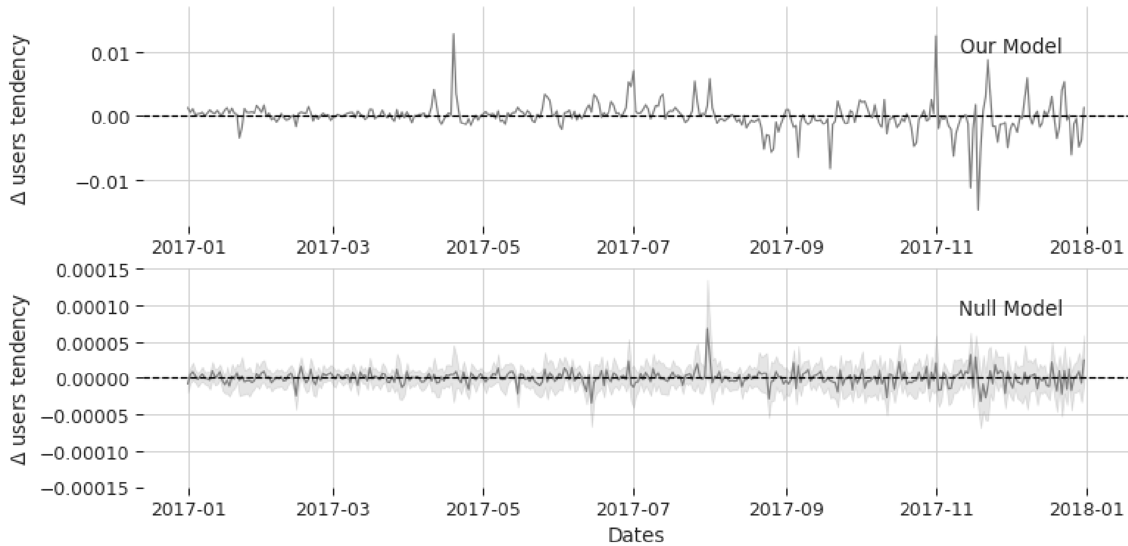


Figure 4.8: Daily difference in the number of users between *empathy* and *threat* groups. Top: our model, and bottom: the null model.

migrant trafficking gang (positive trend spike).

- 06/09: The Minister of Home Affairs, Mario Fernández, faced an appeal at the National Congress after a delay in the Migration Law (negative trend spike).

- 19/09: A sign was installed in Talca (a city on central Chile) urging Haitians to join the Communist Party (negative trend spike).

- 01/11: Director of the Central Public Hospital declared that Joane Florvil (Haitian woman who died after being arrested by Chilean police) had been beaten at the police station (positive trend spike).

- 15/11: Senator Fulvio Rossi from Antofagasta (a city in northern Chile) was stabbed. He stated that "the attacker had a foreign accent and would be a black person" (negative trend spike).

- 18/11: Haitian immigrants attacked police in a commercial neighbourhood in downtown Santiago (negative trend spike).
- 22/11: (1) The Court declared the posthumous innocence of Joane Florvil. (2) Michelle Bachelet recognised not only the heroic act of Richard Joseph (Haitian citizen who rescued a woman who fell from a building) but also a set of positive human values in migration (positive trend spike).

We performed a similar analysis estimating the daily differences in the number of tweets between threat and empathy groups (see Figure 4.9). The results from this analysis display similar patterns of spikes as those found examining differences in user engagement. However, it highlights two key events broadcasted by the national news media: (i) Michelle Bachelet's announcement of a new visa for migrant children and youth through the "Chile welcomes you" program on July, 26th (positive trend spike); and (ii) a confirmed case of an Haitian citizen with leprosy in Valdivia (a city in southern Chile) on July, 31st (negative trend spike).

We validated the differences in both measurements by comparing them with a null model, where attitudes were assigned at random (maintaining the original distribution of attitudes) in 1 K dataset permutations. Figures 4.8 and 4.9 display the null model timeseries. Given that the identified peaks are far from the 95% interval of each null model, the patterns described here are significant.

These results highlight the important role of national news media outlets in shaping the formation and expression of attitudes towards immigration in Chile. All the key events identified above were broadcasted via national television and featured on major national newspapers. As highlighted in intergroup contact and threat theories, the way in which these events are portrayed may contribute to the formation and intensified manifestation of empathetic and threatening attitudes on

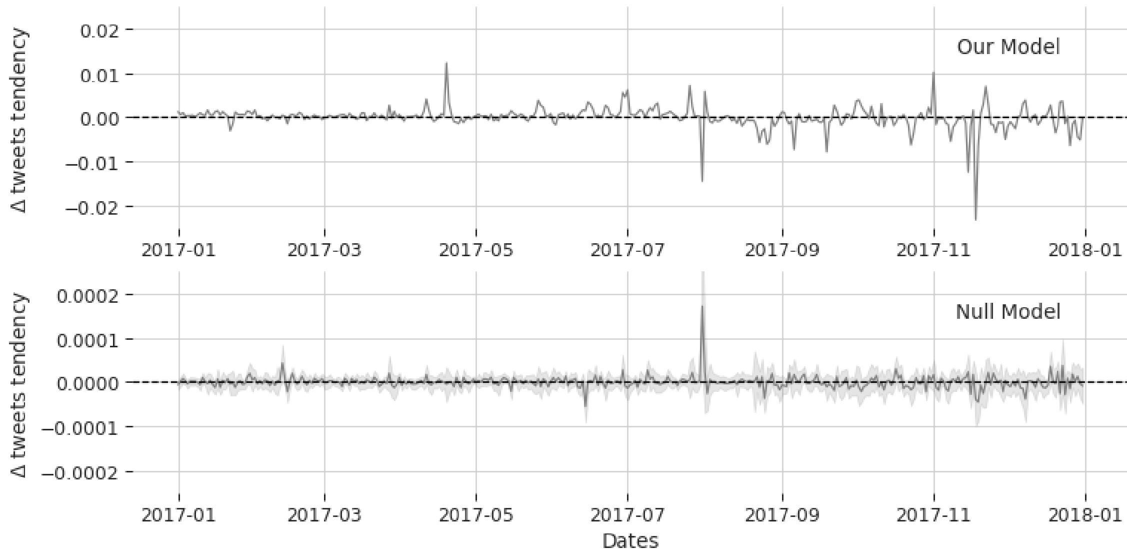


Figure 4.9: Daily difference in the number of tweets between *empathy* and *threat* groups. Top; our model, and bottom; the null model.

social media. For example, highlighted cases of physical attacks by immigrants may foment threat perceptions to public safety and hence invite expressions of negative behaviour towards immigration. The case of the Haitian citizen with leprosy is an interesting example, coinciding with intensification of tweet content generation by *threat* users and exceeding that produced by *empathy* users, despite a greater number of the latter group engaged in the discussion.

In contrast, other events portrayed by the news media were met with empathetic reactions. These include: the creation of a new visa for migrant children and young people, the case of a migrant trafficking gang, of a Haitian citizen who died while being arrested by the police, and immigrants being depicted as victims of violence and discrimination. A prominent example is also the case of a Haitian citizen who rescued a Chilean woman featuring how immigrants can contribute to the local society.

4.6 Discussion

4.6.1 Key Results

Social media is a new, dynamic and open space to express opinions and feelings, collect data and better understand public perception of immigration and offers the opportunity to overcome key limitations of traditional data sources. Social media offers the potential to monitor online public opinion about migration in near-real time at unprecedented spatial and temporal granularity, and understand the actual extent of empathy or negativity in these opinions in raw format. To unlock these potentials, we developed a novel, reproducible and open framework to measure and monitor changes in immigration sentiment. Particularly, the proposed framework enables the classification of users' stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and differences in engagement and content generation at a daily temporal scale.

We presented evidence of the composition of the Chilean immigration sentiment network on Twitter. We found that a larger share (72%) of the user network displays positive and empathetic sentiment towards immigration, and that the proportion of the network associated with negative immigration attitudes was small (25%). However, we also found that users displaying negative anti-immigration profiles tended to produce content at a significantly greater rate, producing up to 50% more content per user than users displaying positive immigration profiles. We find interesting the apparent contrast between public surveys held in Chile and our findings: reportedly, a majority of Chilean nationals agree with the following statements: "the country should take more drastic measures to exclude illegal immigrants." However, the survey is measuring something different: agreement with a statement regarding a specific group of people rather than the expression of an attitude toward a phenomenon. Furthermore, the population under study is different, as the survey samples the population aiming at national representativeness. Thus, a future line of work should be the measurement of representativeness of the Twitter population, and the definition of a methodology to compare survey results with results from our work.

We present a characterisation of the feature and psycholinguistic content of positive and negative immigration user profiles. Users with a pro-immigration profile often use neutral terms recognising migration as a process, and emotional content relating to social, positive emotions, communication, school, humans and family. By contrast, users with an anti-immigration profile tend to use terms denoting a difference between immigrants (*them*) and the native population (*us*), and emotional content relating to motion, other people, present, negative emotions, death, inhibition, anger, money, anxiety and work. We also found evidence of strong alignment between users with anti-immigration views and conservative political ideologies. Our temporal analysis also revealed pronounced peaks in daily user engagement and content generation activity in response to key migration-related events, particularly events featured in news media.

4.6.2 Interpretation

Our findings of a dominant base of users with pro-immigration profiles are consistent with existing prevalent trends in most of the world's regions [55]. Our findings also suggest that although the user base with anti-immigration profiles may be small, it produces and disseminates content at a significantly faster rate. Similar to the effect generated by fake news, the degree of novelty of anti-immigration content and resulting emotional reactions may be the cause of its rapid spread and generation [151].

We also showed that empathetic user profiles were linked to positive emotions, inspiring respect and unity largely in support of immigrants' human and civil rights. Anti-immigration user profiles were associated with negative emotions, calling for stricter immigration laws and claims of migrants "stealing" jobs from locals. We also found evidence of strong alignment between anti-immigration user profiles and conservative political ideologies. This pattern is not specific to Chile, but it is prevalent across industrialised countries [43].

Our analysis revealed high variability in daily user engagement and content generation activity as a result of key migration-related events. These events were prominently featured in national news media, highlighting the pivotal role that news media outlets may play in shaping the formation and expression of attitudes towards immigration in Chile. This calls for a careful approach in the way in which news media outlets portray news items involving immigrants.

4.6.3 Limitations and Future Work

There are two key aspects that need further exploration and that limit the scope of our results: the dynamic analysis of LIWC categories associated with attitudes, and the representativeness of Twitter. In terms of dynamics, it would be interesting to study the temporal distribution of psycholinguistic categories within empathic and threatened attitudes. This would enable quantifying the potential influence of news events on attitudes. This would provide a way to measure the effect of events and their depictions and narratives on how people feel with respect to migration.

In terms of representativeness, we acknowledge that Twitter is a biased sample of the population [5]. However, Twitter is one of the most widely used applications in Chile [73], and it reflects some of its cultural aspects, such as the centralisation of the country [74]. Furthermore, a Twitter-based analysis of the abortion discussion in Chile was found to present equal insights as those from the main national survey that covered the issue [71], hinting that there are social insights that can be derived from the platform. Thus, we propose that this work provides insights with respect to the discussion, although the representativeness of such insights is yet to be determined.

One line of work we sought to explore was to conduct a spatial analysis of attitudes, to understand the relationship between attitudes and the actual presence of immigrants in a place. This would provide a way to measure real and imagined threat attitudes [96], link virtual and physical places of expression and coexistence in a single analytical framework, and allow us to identify socio-demographic characteristics associated with the various inferred attitudes. However, as documented in the literature [132], the spatial representation of Twitter data is limited. Less than three per cent of tweets are geolocated [1]. We recognise that addressing these biases is an active area of research, and case-specific weighting schemes have been proposed to ensure the statistical and spatial representativeness of social media data [81]. However, developing weighting schemes requires knowledge of social media users' profiles. While this information can be obtained with some level of accuracy from Facebook, access to Twitter users' personal attributes is very limited. Therefore, we cannot guarantee that our results represent the general population.

4.7 Conclusions

In this study, we present an analytical framework for monitoring attitudes towards immigration. Specifically the proposed framework enables the classification of users' stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and temporal stance trends.

We applied the proposed framework to 2017 Twitter data from Chile, to capture changes in the virtual public discussion about migration during a period of a surge in immigration. We presented evidence of positive *empathetic* attitudes being expressed by a broad group of users, representing expressions of support for the immigrant community. Particularly these supportive expressions relate to calls for respect, dignity and treatment of immigrants' human and civil rights. Conversely, we provided evidence revealing that negative *threatening* attitudes towards immigration emerge from a reduced number of users, and that these attitudes are prevalent in discussions calling for stricter migrant regulation and concerns about labour competition. We also showed that negative attitudes are more commonly manifested and tend to intensify during instances of negative portrayals of immigrants. These results suggest that media news outlets play a critical role in the spread of negative representations of immigrants, and highlight the need for a more careful approach in the way in which events involving migrants are communicated. Media news outlets should consider the potential impact of misinformation fuelling misconceptions and prejudiced behaviour against immigrants. More broadly, our results demonstrate the need for a systematic approach to monitor immigration sentiment and identify shifts in attitudes towards immigrants. Such approach can enable rapid and effective mitigation plans to address misconceptions and prejudice comments against immigrants and to cushion the long-term formation of negative migration attitudes and their detrimental impacts on national social cohesion.

Bibliography

[1] Twitter (2021b). Tutorials: Tweet geospatialmetadata. https://developer.twitter.com/en/docs/tutorials/tweet-geo-metadata, 2021.

[2] Yann Algan, Christian Dustmann, Albrecht Glitz, and Alan Manning. The economic situation of first and second-generation immigrants in france, germany and the united kingdom, 2010.

[3] Gordon Willard Allport, Kenneth Clark, and Thomas Pettigrew. The nature of prejudice. 1954.

[4] Yair Amichai-Hamburger and Katelyn YA McKenna. The contact hypothesis reconsidered: Interacting via the internet. *Journal of Computer-Mediated Communication*, 11(3):825–843, 2006.

[5] Ricardo Baeza-Yates. Bias on the web. *Communications of the ACM*, 61(6):54–61, 2018.

[6] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval: the concepts and technology behind search, 2nd. Edition*. Addison-Wesley, Pearson, 2011.

[7] Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout,

and Alexander Volfovsky. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221, 2018.

[8] Pablo Barberá. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1):76–91, 2015.

[9] Fiona Kate Barlow, Winnifred R Louis, and Miles Hewstone. Rejected! cognitions of rejection and intergroup anxiety as mediators of the impact of cross-group friendships on prejudice. *British Journal of Social Psychology*, 48(3):389–405, 2009.

[10] Valerio Basile, Cristina Bosco, Elisabetta Fersini, Nozza Debora, Viviana Patti, Francisco Manuel Rangel Pardo, Paolo Rosso, Manuela Sanguinetti, et al. Semeval-2019 task 5: Multilingual detection of hate speech against immigrants and women in twitter. In *13th International Workshop on Semantic Evaluation*, pages 54–63. Association for Computational Linguistics, 2019.

[11] Catalina. Batarce. La tercera. ciudadano haitiano en valdivia es sospechoso de padecer lepra. http://www2.latercera.com/noticia/haitiano-se-convierte-primer-caso-lepra-chile-continental/, 2018.

[12] Barbara E Berger, Carol Estwing Ferrans, and Felissa R Lashley. Measuring stigma in people with hiv: Psychometric assessment of the hiv stigma scale¶. *Research in nursing & health*, 24(6):518–529, 2001.

[13] John W Berry. Acculturation: Living successfully in two cultures. *International journal of intercultural relations*, 29(6):697–712, 2005.

[14] Marianne Bertrand and Esther Duflo. Field experiments on discrimination. *Handbook of Economic Field Experiments*, 1:309–393, 2017.

[15] David M Blei. Probabilistic topic models. *Communications of the ACM*, 55(4):77–84, 2012.

[16] Scott Blinder and William Allen. Uk public opinion toward immigration: Overall attitudes and level of concern. *The Migration Observatory*, 28, 2016.

[17] George J Borjas. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374, 2003.

[18] George J Borjas, Richard B Freeman, Lawrence F Katz, John DiNardo, and John M Abowd. How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, 1997(1):1–90, 1997.

[19] Cristina Bosco, Viviana Patti, Marcello Bogetti, Michelangelo Conoscenti, Giancarlo Francesco Ruffo, Rossano Schifanella, and Marco Stranisci. Tools and resources for detecting hate and prejudice against immigrants in social media. In *SYMPOSIUM III. SOCIAL INTERACTIONS IN COMPLEX INTELLIGENT SYSTEMS (SICIS) at AISB 2017*, pages 79–84. AISB, 2017.

[20] Alexandre Bovet and Hernán A Makse. Influence of fake news in twitter during the 2016 us presidential election. *Nature Communications*, 10(1):1–14, 2019.

[21] Kendrick T Brown, Tony N Brown, James S Jackson, Robert M Sellers, and Warde J Manuel. Teammates on and off the field? contact with black

teammates and the racial attitudes of white student athletes 1. *Journal of applied social psychology*, 33(7):1379–1403, 2003.

[22] John Bryden, Sebastian Funk, and Vincent AA Jansen. Word usage mirrors community structure in the online social network twitter. *EPJ Data Science*, 2(1):1–9, 2013.

[23] Konrad B Burchardi, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry. Immigration, innovation, and growth. Technical report, National Bureau of Economic Research, 2020.

[24] Peter Burns and James G Gimpel. Economic insecurity, prejudicial stereotypes, and public opinion on immigration policy. *Political science quarterly*, 115(2):201–225, 2000.

[25] Carlos Arcila Calderón, Gonzalo de la Vega, and David Blanco Herrero. Topic modeling and characterization of hate speech against immigrants on twitter around the emergence of a far-right party in spain. *Social Sciences*, 9(11):188, 2020.

[26] David Card. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64, 2001.

[27] Héctor Carvacho. Ideological configurations and prediction of attitudes toward immigrants in chile and germany. *International Journal of Conflict and Violence (IJCV)*, 4(2):220–233, 2010.

[28] Public Studies Center. Estudio nacional de opinión pública n°79, abril-mayo 2017. https://www.cepchile.cl/cep/encuestas-cep/encuestas-2010-2019/estudio-nacional-de-opinion-publica-abril-mayo-2017, 2017.

[29] European Political Strategy Centre. European commission. 10 trends shaping migration. https://op.europa.eu/s/oq7V, 2019.

[30] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd International Conference on Knowledge Discovery and Data Mining*, pages 785–794, 2016.

[31] Pauline Hope Cheong, Rosalind Edwards, Harry Goulbourne, and John Solomos. Immigration, social cohesion and social capital: A critical review. *Critical Social Policy*, 27(1):24–49, 2007.

[32] Melanie Coates. Covid-19 and the rise of racism. *Bmj*, 369, 2020.

[33] Gloria Comandini and Viviana Patti. An impossible dialogue! nominal utterances and populist rhetoric in an italian twitter corpus of hate speech against immigrants. In *Third Workshop on Abusive Language Online*, pages 163–171. Association for Computational Linguistics, 2019.

[34] Michael Conover, Jacob Ratkiewicz, Matthew R Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. Political polarization on twitter. *Proc. of ICWSM*, 133:89–96, 2011.

[35] Michael D Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. Predicting the political alignment of twitter users. In *2011 IEEE Third International Conference on Privacy, Security, Risk and*

Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), pages 192–199, Boston, MA, USA, 2011. IEEE.

[36] Glen Coppersmith, Mark Dredze, and Craig Harman. Quantifying mental health signals in twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60, 2014.

[37] Richard J Crisp and Rhiannon N Turner. Can imagined interactions produce positive perceptions?: Reducing prejudice through simulated social contact. *American psychologist*, 64(4):231, 2009.

[38] Lucía Dammert and Matthias Erlandsen. Migración, miedos y medios en la elección presidencial en chile (2017). *Revista CS*, pages 43–76, 2020.

[39] Kareem Darwish, Walid Magdy, Afshin Rahimi, Timothy Baldwin, Norah Abokhodair, et al. Predicting online islamophobic behavior after# parisattacks. *The Journal of Web Science*, 4(3):34–52, 2018.

[40] Mariam Davtyan, Brandon Brown, and Morenike Oluwatoyin Folayan. Addressing ebola-related stigma: lessons learned from hiv/aids. *Global health action*, 7(1):26058, 2014.

[41] Munmun De Choudhury, Scott Counts, Eric J Horvitz, and Aaron Hoff. Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 626–638. ACM, 2014.

[42] James Dennison and Lenka Dražanová. Public attitudes on migration: Rethinking how people perceive migration: An analysis of existing opinion polls

in the euro-mediterranean region. Technical report, European University Institute, 2018.

[43] James Dennison and Andrew Geddes. A rising tide? the salience of immigration and the rise of anti-immigration political parties in western europe. *The Political Quarterly*, 90(1):107–116, 2019.

[44] John F Dovidio, Anja Eller, and Miles Hewstone. Improving intergroup relations through direct, extended and other forms of indirect contact. *Group processes & intergroup relations*, 14(2):147–160, 2011.

[45] Stefan Eriksson. Utrikes födda på den svenska arbetsmarknaden. 2010.

[46] European Social Survey (ESS). Attitudes towards immigration in europe: myths and realities. https://www.europeansocialsurvey.org/docs/findings/IE_Handout_FINAL.pdf, 2017.

[47] Victoria M Esses, John F Dovidio, Lynne M Jackson, and Tamara L Armstrong. The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social issues*, 57(3):389–412, 2001.

[48] P. Darlington. F. Rowe. Quantifying and understanding the extent of residential segregation of recent immigrants in chile. Technical report, prepint, 2021.

[49] Francesco Fasani and Jacopo Mazza. Immigrant key workers: their contribution to europe's covid-19 response. 2020.

[50] Jason Faulkner, Mark Schaller, Justin H Park, and Lesley A Duncan. Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes. *Group Processes & Intergroup Relations*, 7(4):333–353, 2004.

[51] Organisation for Economic Co-operation and Development. *What is the impact of the COVID-19 pandemic on immigrants and their children?* OECD Publishing, 2020.

[52] Economic Commission for Latin America and the Caribbean (ECLAC). Demographic observatory of latin america 2018 : International migration. https://www.oecd.org/acerca/miembros-y-socios/, 2018.

[53] Global Compact for Migration. Global compact for safe, orderly and regular migration. https://www.ohchr.org/en/migration/global-compact-safe-orderly-and-regular-migration-gcm, 2018.

[54] International Organization for Migration. World migration report 2022. https://worldmigrationreport.iom.int/wmr-2022-interactive/, 2022.

[55] International Organization for Migration (IOM). How the world views migration, geneva. https://publications.iom.int/system/files/how_the_world_gallup.pdf, 2015.

[56] International Organization for Migration (IOM). Data bulletin series. informing the implementation of the global compact for migration. https://publications.iom.int/system/files/pdf/gmdacbulletins.pdf, 2018.

[57] Paula Fortuna and Sérgio Nunes. A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, 51(4):1–30, 2018.

[58] Yerka Freire-Vidal and Eduardo Graells-Garrido. Characterization of local attitudes toward immigration using social media. In *Companion Proceedings of The 2019 World Wide Web Conference*, pages 783–790, 2019.

[59] Yerka Freire-Vidal, Eduardo Graells-Garrido, and Francisco Rowe. A framework to understand attitudes towards immigration through twitter. *Applied Sciences*, 11(20):9689, 2021.

[60] Yoav Freund, Robert Schapire, and Naoki Abe. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612, 1999.

[61] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[62] Ryan J Gallagher, Morgan R Frank, Lewis Mitchell, Aaron J Schwartz, Andrew J Reagan, Christopher M Danforth, and Peter Sheridan Dodds. Generalized word shift graphs: A method for visualizing and explaining pairwise comparisons between texts. *arXiv preprint arXiv:2008.02250*, 2020.

[63] Ruth Garcia-Gavilanes, Daniele Quercia, and Alejandro Jaimes. Cultural dimensions in twitter: Time, individualism and power. *Proc. of ICWSM*, 13, 2013.

[64] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. The effect of collective attention on controversial debates on social media. In *Proceedings of the 2017 ACM on Web Science Conference*, pages 43–52, 2017.

[65] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, pages 913–922, Lyon, France, 2018. International World Wide Web Conferences Steering Committee.

[66] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1(1):3, 2018.

[67] Julia Gelatt. Immigrant workers: Vital to the us covid-19 response, disproportionately vulnerable. *Migration Policy Institute*, 26, 2020.

[68] Scott A Golder and Michael W Macy. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051):1878–1881, 2011.

[69] Roberto González-Ibánez, Smaranda Muresan, and Nina Wacholder. Identifying sarcasm in twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*, pages 581–586. Association for Computational Linguistics, 2011.

[70] Eduardo Graells-Garrido and Ricardo Baeza-Yates. Bots don't vote, but they surely bother! a study of anomalous accounts in a national referendum. In *14th ACM Web Science Conference 2022*, pages 302–306, 2022.

[71] Eduardo Graells-Garrido, Ricardo Baeza-Yates, and Mounia Lalmas. How representative is an abortion debate on twitter? In *Proceedings of the 10th ACM Conference on Web Science*, pages 133–134, 2019.

[72] Eduardo Graells-Garrido, Ricardo Baeza-Yates, and Mounia Lalmas. Every colour you are: Stance prediction and turnaround in controversial issues. In *12th ACM Conference on Web Science*, WebSci '20, page 174–183, New York, NY, USA, 2020. Association for Computing Machinery.

[73] Eduardo Graells-Garrido, Diego Caro, Omar Miranda, Rossano Schifanella, and Oscar F Peredo. The www (and an h) of mobile application usage in the city: The what, where, when, and how. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 1221–1229. International World Wide Web Conferences Steering Committee, 2018.

[74] Eduardo Graells-Garrido and Mounia Lalmas. Balancing diversity to counter-measure geographical centralization in microblogging platforms. In *Proceedings of the 25th ACM conference on Hypertext and social media*, pages 231–236. ACM, 2014.

[75] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. Finding intermediary topics between people of opposing views: a case study. CEUR, Santiago, Chile.

[76] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. Data portraits and intermediary topics: Encouraging exploration of politically diverse profiles. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pages 228–240, 2016.

[77] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. Encouraging diversity-and representation-awareness in geographically centralized content. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pages 7–18. ACM, 2016.

[78] Thomas Greven. The rise of right-wing populism in europe and the united states. *A Comparative Perspective [La emergencia del populismo de derechas en Europa y Estados Unidos. Una perspectiva comparada]. Friedrich Ebert Foundation, Washington DC Office*, pages 1–8, 2016.

[79] Alexis Grigorieff, Christopher Roth, and Diego Ubfal. Does information change attitudes toward immigrants? *Demography*, 57(3):1117–1143, 2020.

[80] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. Fake news on twitter during the 2016 us presidential election. *Science*, 363(6425):374–378, 2019.

[81] André Grow, Daniela Perrotta, Emanuele Del Fava, Jorge Cimentada, Francesco Rampazzo, Sofia Gil-Clavel, Emilio Zagheni, René D Flores, Ilana Ventura, Ingmar Weber, et al. How reliable is facebook's advertising data for use in social science research? insights from a cross-national online survey. Technical report, Max Planck Institute for Demographic Research, Rostock, Germany, 2021.

[82] Shang E Ha. The consequences of multiracial contexts on public attitudes toward immigration. *Political Research Quarterly*, 63(1):29–42, 2010.

[83] Jens Hainmueller and Michael J Hiscox. Educated preferences: Explaining attitudes toward immigration in europe. *International organization*, 61(2):399–442, 2007.

[84] Jens Hainmueller and Michael J Hiscox. Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American political science review*, 104(1):61–84, 2010.

[85] Jens Hainmueller and Daniel J Hopkins. Public attitudes toward immigration. *Annual review of political science*, 17:225–249, 2014.

[86] Gordon H Hanson, Kenneth Scheve, and Matthew J Slaughter. Public finance and individual preferences over globalization strategies. *Economics & Politics*, 19(1):1–33, 2007.

[87] GACCT Harman and Mark H Dredze. Measuring post traumatic stress disorder in twitter. *In ICWSM*, 2014.

[88] Gregory M Herek. The instrumentality of attitudes: Toward a neofunctional theory. *Journal of Social Issues*, 42(2):99–114, 1986.

[89] Gregory M Herek and John P Capitanio. “some of my best friends” intergroup contact, concealable stigma, and heterosexuals’ attitudes toward gay men and lesbians. *Personality and social psychology bulletin*, 22(4):412–424, 1996.

[90] Danny Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on visualization and computer graphics*, 12(5):741–748, 2006.

[91] Daniel J Hopkins. Politicized places: Explaining where and when immigrants provoke local opposition. *American political science review*, 104(1):40–60, 2010.

[92] Daniel J Hruschka and Joseph Henrich. Institutions, parasites and the persistence of in-group preferences. *PloS one*, 8(5):e63642, 2013.

[93] INSTITUTO NACIONAL DE ESTADÍSTICAS (INE). Informe de resultados de la estimación de personas extranjeras residentes en chile al 31 de diciembre de 2021. https://www.ine.gob.cl/docs/default-source/demografia-y-migracion/publicaciones-y-anuarios/migración-internacional/estimación-población-extranjera-en-chile-2018/estimación-población-extranjera-en-chile-2021-resultados.pdf?sfvrsn=d4fd5706_6, 2022.

[94] National Statistics Institute (INE). Síntesis resultados censo 2017. https://www.censo2017.cl/descargas/home/sintesis-de-resultados-censo2017.pdf, 2018.

[95] Seth K Jolly and Gerald M DiGiusto. Xenophobia and immigrant contact: French public attitudes toward immigration. *The Social Science Journal*, 51(3):464–473, 2014.

[96] Jeffrey S Kopstein and Jason Wittenberg. Does familiarity breed contempt? inter-ethnic contact and support for illiberal parties. *The Journal of Politics*, 71(2):414–428, 2009.

[97] Adam DI Kramer. An unobtrusive behavioral model of gross national happiness. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 287–290. ACM, 2010.

[98] Onur Kucuktunc, B Barla Cambazoglu, Ingmar Weber, and Hakan Ferhatosmanoglu. A large-scale sentiment analysis for yahoo! answers. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 633–642. ACM, 2012.

[99] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, pages 591–600, 2010.

[100] Fabio Lamanna, Maxime Lenormand, María Henar Salas-Olmedo, Gustavo Romanillos, Bruno Gonçalves, and José J Ramasco. Immigrant community integration in world cities. *PloS one*, 13(3):e0191612, 2018.

[101] Duncan Lawrence. Crossing the cordillera: immigrant attributes and chilean attitudes. *Latin American Research Review*, 50(4):154–177, 2015.

[102] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788, 1999.

[103] Debra Lieberman and Carlton Patrick. Are the behavioral immune system and pathogen disgust identical? *Evolutionary Behavioral Sciences*, 8(4):244, 2014.

[104] Bruce G Link and Jo C Phelan. Stigma and its public health implications. *The Lancet*, 367(9509):528–529, 2006.

[105] Haokai Lu, James Caverlee, and Wei Niu. Biaswatch: A lightweight system for discovering and tracking topic-sensitive opinion bias in social media. In *Proceedings of the 24th ACM International on Conference on Information*

and Knowledge Management, pages 213–222, Melbourne, Australia, 2015. ACM, ACM.

[106] Yao Lu, Peng Zhang, Yanan Cao, Yue Hu, and Li Guo. On the frequency distribution of retweets. *Procedia Computer Science*, 31:747–753, 2014.

[107] Kelsey MacMillan and James D Wilson. Topic supervised non-negative matrix factorization. *arXiv preprint arXiv:1706.05084*, 2017.

[108] Ajaykumar Manivannan, W Quin Yow, Roland Bouffanais, and Alain Barrat. Are the different layers of a social network conveying the same information? *EPJ Data Science*, 7(1):34, 2018.

[109] Burt L Monroe, Michael P Colaresi, and Kevin M Quinn. Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403, 2008.

[110] Chad R Mortensen, D Vaughn Becker, Joshua M Ackerman, Steven L Neuberg, and Douglas T Kenrick. Infection breeds reticence: The effects of disease salience on self-perceptions of personality and behavioral avoidance tendencies. *Psychological science*, 21(3):440–447, 2010.

[111] Carlos David Navarrete and Daniel MT Fessler. Disease avoidance and ethnocentrism: The effects of disease vulnerability and disgust sensitivity on intergroup attitudes. *Evolution and Human Behavior*, 27(4):270–282, 2006.

[112] Todd D Nelson. *Handbook of prejudice, stereotyping, and discrimination*. Psychology Press, 2009.

[113] Adina Nerghes and Ju-Sung Lee. The refugee/migrant crisis dichotomy on twitter: A network and sentiment perspective. In *Proceedings of the 10th ACM Conference on Web Science*, pages 271–280, 2018.

[114] Mark EJ Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, 2003.

[115] Josef Novotny and Filip Polonsky. The level of knowledge about islam and perception of islam among czech and slovak university students: does ignorance determine subjective attitudes? *Sociologia*, 43(6):674–696, 2011.

[116] Kevin H O'rourke and Richard Sinnott. The determinants of individual attitudes towards immigration. *European journal of political economy*, 22(4):838–861, 2006.

[117] Elizabeth Levy Paluck, Seth A Green, and Donald P Green. The contact hypothesis re-evaluated. *Behavioural Public Policy*, 3(2):129–158, 2019.

[118] Bo Pang, Lillian Lee, et al. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135, 2008.

[119] James W Pennebaker, Martha E Francis, and Roger J Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[120] James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1):547–577, 2003.

[121] Rinus Penninx, Dimitrina Spencer, Nicholas Van Hear, et al. Migration and integration in europe: The state of research. *Swindon, UK: Economic and Social Research Council*, 2008.

[122] Thomas F Pettigrew and Linda R Tropp. A meta-analytic test of intergroup contact theory. *Journal of personality and social psychology*, 90(5):751, 2006.

[123] Thomas F Pettigrew and Linda R Tropp. How does intergroup contact reduce prejudice? meta-analytic tests of three mediators. *European Journal of Social Psychology*, 38(6):922–934, 2008.

[124] Simon Porcher and Thomas Renault. Social distancing beliefs and human mobility: Evidence from twitter. *Plos one*, 16(3):e0246949, 2021.

[125] Panu Poutvaara and Max Friedrich Steinhardt. Bitterness in life and attitudes towards immigration. *European Journal of Political Economy*, 55:471–490, 2018.

[126] Naomi Priest, Yin Paradies, Angeline Ferdinand, Lobna Rouhani, and Margaret Kelaher. Patterns of intergroup contact in public spaces: Micro-ecology of segregation in australian communities. *Societies*, 4(1):30–44, 2014.

[127] Daniele Quercia, Jonathan Ellis, Licia Capra, and Jon Crowcroft. Tracking gross community happiness from tweets. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pages 965–968. ACM, 2012.

[128] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. Our twitter profiles, our selves: Predicting personality with twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 180–185. IEEE, 2011.

[129] Nairán Ramírez-Esparza, James W Pennebaker, Florencia Andrea García, and Raquel Suriá. La psicología del uso de las palabras: Un programa de computadora que analiza textos en español. *Revista Mexicana de Psicología*, 24(1):85–99, 2007.

[130] Manoel Ribeiro, Pedro Calais, Yuri Santos, Virgílio Almeida, and Wagner Meira Jr. Characterizing and detecting hateful users on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12, 2018.

[131] Ryan S Ritter, Jesse Lee Preston, and Ivan Hernandez. Happy tweets: Christians are happier, more socially connected, and less analytical than atheists on twitter. *Social Psychological and Personality Science*, 5(2):243–249, 2014.

[132] Francisco Rowe, Michael Mahony, Eduardo Graells-Garrido, Marzia Rango, and Niklas Sievers. Using twitter to track immigration sentiment during early stages of the covid-19 pandemic. *Data & Policy*, 3:e36, 2021.

[133] Francisco Rowe, Michael Mahony, Eduardo Graells-Garrido, Marzia Rango, and Niklas Sievers. Using Twitter to track immigration sentiment during early stages of the COVID-19 pandemic. *Data & Policy (in review)*, 2021.

[134] Manuela Sanguinetti, Fabio Poletto, Cristina Bosco, Viviana Patti, and Marco Stranisci. An italian twitter corpus of hate speech against immigrants. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

[135] Georgia Sarolidou, John Axelsson, Bruce A Kimball, Tina Sundelin, Christina Regenbogen, Johan N Lundström, Mats Lekander, and Mats J Olsson. People expressing olfactory and visual cues of disease are less liked. *Philosophical Transactions of the Royal Society B*, 375(1800):20190272, 2020.

[136] Mark Schaller. Parasites, behavioral defenses, and the social psychological mechanisms through which cultures are evoked. *Psychological Inquiry*, 17(2):96–101, 2006.

[137] Mark Schaller and Justin H Park. The behavioral immune system (and why it matters). *Current directions in psychological science*, 20(2):99–103, 2011.

[138] Kenneth F Scheve and Matthew J Slaughter. Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, 83(1):133–145, 2001.

[139] M Ángeles Serrano, Alessandro Flammini, and Filippo Menczer. Modeling statistical properties of written text. *PloS one*, 4(4):e5372, 2009.

[140] Paul M Sniderman, Pierangelo Peri, Rui JP de Figueiredo Jr, and Thomas L Piazza. The outsider: Politics and prejudice in italy, 2000.

[141] Cookie White Stephan and Walter S Stephan. An integrated threat theory of prejudice. In *Reducing prejudice and discrimination*, pages 33–56. Psychology Press, 2013.

[142] Walter G Stephan and Krystina Finlay. The role of empathy in improving intergroup relations. *Journal of Social issues*, 55(4):729–743, 1999.

[143] Walter G Stephan and Cookie White Stephan. Intergroup anxiety. *Journal of social issues*, 41(3):157–175, 1985.

[144] Karolina Sylwester and Matthew Purver. Twitter language use reflects psychological differences between democrats and republicans. *PloS one*, 10(9):e0137422, 2015.

[145] Andranik Tumasjan, Timm O Sprenger, Philipp G Sandner, and Isabell M Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Fourth International AAAI Conference on Weblogs and Social Media*, 2010.

[146] Joshua M Tybur, Debra Lieberman, and Vladas Griskevicius. Microbes, mating, and morality: individual differences in three functional domains of disgust. *Journal of personality and social psychology*, 97(1):103, 2009.

[147] Joshua M Tybur, Debra Lieberman, Robert Kurzban, and Peter DeScioli. Disgust: evolved function and structure. *Psychological review*, 120(1):65, 2013.

[148] United Nations (UN). Universal declaration of human rights. https://www.un.org/en/about-us/universal-declaration-of-human-rights, 1948.

[149] United Nations (UN). United nations. sustainable development goals. https://www.un.org/sustainabledevelopment/sustainable-development-goals/, 2015.

[150] Paula Vasquez-Henriquez, Eduardo Graells-Garrido, and Diego Caro. Tweets on the go: Gender differences in transport perception and its discussion on social media. *Sustainability*, 12(13):5405, 2020.

[151] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[152] Zeerak Waseem and Dirk Hovy. Hateful symbols or hateful people? predictive features for hate speech detection on twitter. In *Proceedings of the NAACL student research workshop*, pages 88–93, 2016.

[153] Fiona A White and Hisham M Abu-Rayya. A dual identity-electronic contact (diec) experiment promoting short-and long-term intergroup harmony. *Journal of Experimental Social Psychology*, 48(3):597–608, 2012.

[154] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 347–354, 2005.

[155] Shaomei Wu, Jake M Hofman, Winter A Mason, and Duncan J Watts. Who says what to whom on twitter. In *Proceedings of The 20th International Conference on World Wide Web*, pages 705–714, 2011.

[156] Qiang Zhang, Shangsong Liang, Aldo Lipani, Zhaochun Ren, and Emine Yilmaz. From stances' imbalance to their hierarchical representation and

detection. In *The World Wide Web Conference*, pages 2323–2332. ACM, 2019.

Chapter 5

APPENDIX

The following are the papers, which are part of this thesis work, and which have already been published.

Characterization of Local Attitudes Toward Immigration Using Social Media

Yerka Freire-Vidal
Social Complexity Research Center
Universidad del Desarrollo
Santiago, Chile
yfreirev@udd.cl

Eduardo Graells-Garrido
Data Science Institute
Universidad del Desarrollo
Santiago, Chile
egraells@udd.cl

ABSTRACT

Migration is a worldwide phenomenon that may generate different reactions in the population. Attitudes vary from those that support multiculturalism and communion between locals and foreigners, to contempt and hatred toward immigrants. Since anti-immigration attitudes are often materialized in acts of violence and discrimination, it is important to identify factors that characterize these attitudes. However, doing so is expensive and impractical, as traditional methods require enormous efforts to collect data. In this paper, we propose to leverage Twitter to characterize local attitudes toward immigration, with a case study on Chile, where immigrant population has drastically increased in recent years. Using semi-supervised topic modeling, we situated 49K users into a spectrum ranging from in-favor to against immigration. We characterized both sides of the spectrum in two aspects: the emotions and lexical categories relevant for each attitude, and the discussion network structure. We found that the discussion is mostly driven by Haitian immigration; that there are temporal trends in tendency and polarity of discussion; and that assortative behavior on the network differs with respect to attitude. These insights may inform policy makers on how people feel with respect to migration, with potential implications on communication of policy and the design of interventions to improve inter-group relations.

ACM Reference Format:
Yerka Freire-Vidal and Eduardo Graells-Garrido. 2019. Characterization of Local Attitudes Toward Immigration Using Social Media. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3308560.3316455

1 INTRODUCTION

Migration is a phenomenon faced by many countries, which brings a variety of effects; both in the population from which it emigrates and in the receiving population. One of the effects that worries many countries is intolerance and hostile attitudes toward immigrants. These attitudes have been the focus of many research studies, some of which are focused on individual-level psychological and socio-economic factors [8, 47], and others on the contact between immigrant population and locals [7, 30, 32]. The main methods used in these studies are based on context specific surveys, which makes replication in others societies or countries difficult. The theories that explain the type of attitudes of locals interacting with immigrants can be summarized in two: the Intergroup Contact Theory [1], and the Integrated Threat Theory [40, 49]. The former states that people support multiculturalism and integration. The latter, that people think that immigrants will bring negative effects for their society, including competition for jobs and public services, worsening of the national economy, increase in crime, and the arrival of diseases. Particularly, the attitudes explained by the threat theory can lead to acts of violence, discrimination, and abuse; thus, it is important to understand what factors enhance such attitudes.

However, measuring attitudes is costly and impractical under dynamic scenarios. The most frequent methods are surveys, which are difficult and costly to implement. In this paper, we propose to make use of the information that people publish in Twitter as a proxy of their attitudes toward immigration. It is common to find reactions and attitudes through posts in these platforms, where people express their ideas and opinions voluntarily. We propose to define a spectrum of attitudes based on the two aforementioned theories, and to classify users and tweets into that spectrum. We do so with a semi-supervised topic modeling technique named Topic-Supervised Non-Negative Matrix Factorization [39]. TS-NMF works in a semi-supervised way because some users can be labeled as belonging to each extreme of the spectrum, something that we do with custom-built lexicons for each theory.

We perform a descriptive case study on the Chilean society, because Chile is one of the countries in which migration has reached unprecedented volume in recent years. The statistics show that immigrant population has increased from 0.8% in 1992 to 4.35% in 2017; and where 66.7% of immigrants declare to have arrived mainly in 2016 [31]. For this, Chileans have developed diverse perceptions regarding the number of immigrants in the country and the phenomenon itself. To measure them with our proposed method, we collected more than 206K tweets that discuss immigration in Chile, written by more than 49K users during the year 2017. After inferring user and tweets positions in the spectrum, we performed lexical and network analysis with respect to the spectrum position. In the lexical analysis, we used the "Linguistic Inquiry and Word Count" (LIWC) lexicon [43], typically employed to characterize cognitive and emotional differences in discourse [11, 17, 26]. To analyze the network structure, we estimated the polarization of the retweet and mention networks between users.

As main results, we observed that most of the discussion toward migration in Chile is targeted at Haitian migration, even though other countries have a larger share of the population. We found

WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308560.3316455

lexical differences in how each attitude discussed migration, and those differences were consistent with theories. For instance, social-related words were correlated with empathetic attitudes, job- and money-related words were correlated with threatening attitudes. In the network, the retweet network was polarized, in coherence as predicted by other studies regarding political discussion [10, 20]. Finally, we notice that the amount and tendency of the tweets (the latter reflects the attitude towards immigration) seems to be influenced by relevant news events on national migration issues. These results can inform public policy designers to improve inter-group relations in the country, as well as increasing the understanding of how people feel regarding an important aspect of globalization.

In summary, the contribution of this paper is two-fold. We proposed a methodology to characterize local attitudes toward migration from tweets. Then, we performed a descriptive case study in Chile using this method, obtaining results that are coherent with social theory, with added depth based on the rich information that can be extracted from Twitter.

This paper is structured as follows. Section 2 discusses the related work. Section 3 describes the social theories that guided our analysis. Section 4 describes the data set we analyzed. Section 5 describes the methodology. Section 6 describes the results of applying the methodology to the data set. Section 7 discusses the implications of our work. Finally, Section 8 states our conclusions.

2 RELATED WORK

Migration is a widely studied topic because there are many issues associated to this phenomenon. Some researchers have focused on studying the economic impacts related to migration [23–25, 47], others on social cohesion [30, 32, 33, 48]. Within these studies, those who have focused on integration [36] and racism/xenophobia stand out [7, 27, 44]. Our work seeks to contribute in the latter area, mainly due to the subject of our case study, Chile, a society that in a short time has faced a massive influx of immigrants. Migration in Chile has been a national issue, causing controversy in presidential elections, news, and municipal institutions. However, measuring attitudes is not a simple problem, nor a solved one. Twitter is currently a platform widely used in studies of human behavior, since it provides a valuable source of data. Studies that have used Twitter have allowed to reveal socio-cultural characteristics of users or societies, including the level of integration of immigrants in a city [36], attitudes in response to triggering events, such as terrorist attacks [13], the influence of culture in personal actions [16], political polarization [16], personality traits [46], and personality differences between democrats and republicans [52].

Given this body of research, we propose that Twitter can be used as a proxy to understand human behavior, in our case, the attitudes of Chileans regarding immigration.

3 SOCIAL THEORIES

The attitudes toward immigration are varied and depend on economic, socio-cultural and psychological factors. In this context, psychology and sociology have defined theories that explain the attitudes exhibited by people, who belong to different groups, when interacting with others: the Intergroup Contact Theory [1], and the Integrated Threat Theory [40, 49]. The attitudes toward immigration are a particular case explained by these theories.

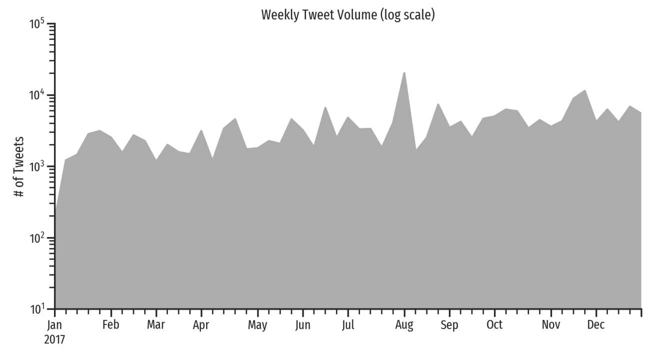


Figure 1: Weekly distribution of tweets about immigration in Chile.

Intergroup Contact Theory. Developed in the book "The Nature of Prejudice" by Gordon W. Allport [1], it postulates that prejudices are reduced when the interaction between different groups meets the following conditions: 1) groups are on equal terms; 2) they have common goals; 3) there is cooperation; and 4) there is support from formal and/or informal institutions. The theory states that intergroup contact reduces the fear and anxiety that exists when people interact with an unknown group [51], and that it promotes empathy and understanding towards the foreign group [50].

This theory has been used to ground several studies: contact between white and black people [7], heterosexuals and homosexuals [27, 28], minority religious groups [42], and locals and immigrants [30]. All these studies conclude that contact improves relationships between groups.

Integrated Threat Theory. In contrast to the Intergroup Contact Theory, the Integrated Threat Theory argues that contact between disparate groups provokes perceptions of threat and contempt [40, 49], for instance, due to competition for work and economic resources [15, 22]. Furthermore, the threat does not have to be real, it can be subjective or fictitious [33].

This theory postulates that, when the interaction conditions are not optimal, the contact between different groups will provoke conflicting and hostile relationships. The concept of "contact" is not limited only to physical contact, it can also be indirect [14], imagined [12], and electronic [2, 53].

Both theories tell us what to search when we look attitudes toward immigration: attitudes motivated by empathy, in favor of immigration; and attitudes motivated by threat, against immigration. As such, we will assume that there are two attitudes, which we label as *empathy* and *threat*.

4 DATA SET DESCRIPTION

In this section we describe our data set of posts from Twitter about migration in Chile.

Twitter is a micro-blogging platform, where users publish tweets (posts) with a maximum of 280 characters. Users may follow others,

Figure 2: Wordcloud of most frequent words in the dataset, after removing stopwords. Color is assigned according to the following categories: words, hashtags, mentions, and URLs.

to see their tweets in their own timelines. Tweets may mention other users, quote other tweets, or retweet another tweet to share it with one's audience. Users can report a screen name, a full name (which can be real or fictitious), a location (real, fictitious, or empty), and a small autobiography, among other attributes. To collect tweets that talk about immigration in Chile we used the Twitter Streaming API using system designed to crawl Chilean tweets [21]. The query parameters were keywords related to immigration (*e.g.*, inmigración, inmigrante, fronteras, racismo, *etc.*), and origin countries with their respective demonyms (*e.g.*, Haití–haitianos/as, Venezuela–venezolanos/as, Perú–peruanos/as, *etc.*). Given how generic some of these keywords are, particularly regarding the context of political issues of neighbouring countries, and the presidential elections held in Chile during November and December, we performed extensive manual clean-up of the data set.

In total, our data set is comprised by 206,353 tweets that discuss immigration in Chile during 2017, written by 49,346 users. Figure 1 shows the weekly volume of tweets. As seen on the figure, the amount of tweets has a sligth positive trend. Two peaks draw our attention: July 31th, when the news reported a case of an Haitian citizen with Leprosy; and November 19th, when an Haitian citizen rescued a woman who fell from the ninth floor of a building.

Regarding content, Figure 2 shows the most frequent words, after removing stopwords and accents. One can see that words such as Haití and Haitianos are more relevant than other countries name or demonyms, despite the fact that the largest immigrant population comes from Perú, Colombia and Venezuela [31]. Also, Santiago and Antofagasta are two frequent keywords, the two cities with the largest immigrant population [31]. Other relevant words that appear are: "gobierno" (government), "carabineros" (policemen), "Piñera" (current president, and presidential candidate in 2017) and "proyecto" (project); possibly because during the year an immigration reform was being discussed.

We explored the data set to seek for words, phrases, and hashtags that could be mapped to the *empathy* and *threat* attitudes. In *empathy* we chose terms that indicated that immigrants are welcome and will be received in equal conditions (*e.g.*, "we are all immigrants"). In *threat* we chose terms and words that showed that immigrants are not welcome and qualified them negatively (*e.g.*, "illegal immigrants"). Table 1 shows some examples of the the

Table 1: Examples of training terms for each attitude.

Attitude	Training Terms
Empathy	#todossomosmigrantes, #stopxenophobia, #chilesinbarreras, #chileterecibe, #bienvenidosmigrantes, @oimchile, bienvenidos a chile, #derribandomuros, @sjmchile, …
Threat	#vendepatria, #nomasinmigrantes, #nomasilegales, #inmigrantesilegales, inmigrantes delincuentes, inmigracion descontrolada, indeseables, …

terms we associated to both attitudes. These labeled terms are not necessarily frequent, however, the methodology that we describe in the next sections allows to propagate these labels through a topic model.

5 METHODOLOGY

In this section we describe how to characterize users and tweets according to their attitude toward immigration. We define how to apply machine learning techniques to user profiles to derive user-attitude and term-attitude associations. Then, we define how to characterize attitudes from sentiment, lexical and network perspectives.

5.1 Attitudes and Topic Modeling

Topic models are a family of techniques used to discover the underlying semantic structure of a corpus by identifying and quantifying the importance of representative themes in all documents [6]. Topic models assume that each text document is generated by a set of topics which have a determined distribution. At the same time, each topic is defined by a set of words, which also have a particular distribution for each topic.

A popular topic modeling technique is Non-negative Matrix Factorization [38]. NMF works by constructing a k-rank factorization of a positive document-term matrix V into $W \times H$. Matrices W and H are estimated by minimizing the following objective function:

$$D_{NMF}(W, H) = \| V - W \times H \|_F^2, \quad W, H \geq 0, \tag{1}$$

where $\| \cdot \|_F$ is the Frobenius norm. In topic modeling, k, W and H have a special interpretation: k is the number of topics, W_{ij} quantifies the relevance of topic j in document i, and H_{ij} quantifies the relevance of term j in topic i.

Typical topic modeling applications select different numbers of k based on metrics such as perplexity. However, the meaning of topics is not always interpretable, as the factorization may follow latent patterns not necessarily aligned with human expectations. Based on the social theories described in Section 3, we propose to guide the learning procedure to seek for two topics: one that represents *empathy*, and another that represents *threat*. In such cases, supervised methods could be employed, however, these methods require a fully labeled data set, not available in our case. Since it is possible to map specific terms (words, phrases, hashtags, URLs, *etc.*) into these two topics, we propose to use a semi-supervised version of NMF known as Topic-Supervised NMF [39]. TS-NMF defines the minimization problem as follows:

$$D_{TS}(W,H) = \| V - (W \circ L)H \|_F^2, \quad W, H \geq 0, \tag{2}$$

where $\circ$ is the Haddamard product operator, and L is a supervision matrix, defined as $L_{ij} = 1$ if topic j contributes to the document i, and $L_{ij} = 0$ if the topic j does not contribute to the document i. Thus, TS-NMF allows to provide examples of documents labeled with known topics, and to restrict the latent representation of the corpus to align with the labeled examples.

In our context, we work with user profiles, *i.e.*, the concatenation of tweets by a single user is one document. As terms we consider hashtags, mentions, URLs, and n-grams with n up to four. This allows us to define how specific phrases are mapped to each topic. The user corpus is represented as a document-term matrix D weighted with TF-IDF [4], and then row-normalized with L2 norm. To label users in the supervision matrix, we construct a list of seed terms for each theory. Then, for each row in D we estimate a preliminary attitude score for each topic, by adding the values of the cells of the corresponding seed terms. All users with a score above a certain threshold are labeled with the corresponding topic. In our experiments, we defined a threshold of 0.25, implying that only users who strongly used the seed terms of each topic were labeled.

As result, we obtain $D = U \times T$, where the rank of U and T is two. In our context, each topic is an attitude, the matrix U contains the user-attitude associations, and the matrix T contains the term-attitude associations (transposed). We interpret these associations as probabilities.

5.2 Attitude Tendency and Polarity

To characterize attitudes, we calculate two metrics common in the sentiment analysis literature to measure the leaning and amount of sentiment: tendency and polarity [35]. Tendency is defined as:

$$\text{tendency}(u) = P(\text{empathy} \mid u) - P(\text{threat} \mid u), \tag{3}$$

where, $P(\text{attitude} \mid u)$ is the association between user u and the corresponding attitude. Note that the definition is analog for terms. For tweets, tendency is defined as:

$$\text{tendency}(\text{tweet}) = \sum_{\text{term} \in \text{tweet}} \text{tendency}(\text{term}). \tag{4}$$

Note that tendency values close to zero do not imply a neutral attitude, as there could be non-zero contributions in both topics. To clarify this fact, we consider attitude polarity as the amount of associations to both attitudes, defined for users as:

$$\text{polarity}(u) = P(\text{empathy} \mid u) + P(\text{threat} \mid u). \tag{5}$$

The definition for terms is analog. For tweets, polarity is defined as:

$$\text{polarity}(\text{tweet}) = \sum_{\text{term} \in \text{tweet}} \text{polarity}(\text{term}). \tag{6}$$

In this way, tendency will allow us to group users/tweets (according to their attitude), while polarity will allow us to measure the intensity of the discussion (how polarized is the attitude).

5.3 Lexical Characterization

The previous metrics give an overview of user and tweet attitudes. The next step is to characterize grouped tweets belonging to each attitude according to their tendency. To do so, we use a psycholinguistic lexicon named "Linguistic Inquiry and Word Count" [43]. LIWC is a lexicon used to study emotional, cognitive and structural components contained in a text. In its Spanish version, it contains 7,515 words classified in one or more of 72 categories. Categories are classified into four dimensions: 1) standard linguistic processes (*e.g.*, articles, prepositions, pronouns, *etc.*); 2) psychological processes (*e.g.*, positive and negative emotions); 3) relativity (*e.g.*, time, verb tense, motion, space); and 4) personal matters (*e.g.*, sex, death, home, occupation, *etc.*). LIWC categories are organized hierarchically, for instance, all words related to the category *anger* are also organized in the categories of *negative emotions* or *affect* words.

We seek to estimate the association of tweets by tendency groups to LIWC categories. After classifying tweets into groups, we estimate how associated the words in LIWC are to each group. Note that specific events may entice a more active discussion by either group, increasing the amount of tweets, thus, we need a way to control the association with these activity patterns. In previous work, this has been done to estimate gross community metrics with z-scores [34, 45]. In our case, the definition is as follows:

$$Z_{lt'} = \frac{P_{lt'} - \mu_l}{\sigma_l}, \tag{7}$$

where, $Z_{lt'}$ is the association of LIWC category l with the tendency t', $P_{lt'}$ is the mean of fraction of words in l in each tweet with tendency t', μ_l is the mean of fraction of words in l in all tweets, and σ_l is the standard deviation of the fraction of words in l in all tweets. Hence, this relative metric allows us to compare behavior between groups, by controlling for external variability.

5.4 Network Assortativity

The previous definitions capture the behavior in expression, however, the social aspect of Twitter allows to also capture network behavior. We focus on two different networks: the mention network, related to discussion, and the retweet network, related to information diffusion. In both networks, node are users, and links are weighted relations between users. Each node has as attributes its associations to each attitude. In the mention network, a directed link between users u_1 and u_2 exists if u_1 mentions u_2 in one or more tweets. The link weight is the number of times this happens. In the retweet network, a directed link between users u_1 and u_2 exists if u_1 republishes content by u_2. The link weight is the number of times that one user retweets another. These kind of networks are commonly analyzed to understand polarization [10]. To be able to analyze connectivity, we will focus on the Largest Strongly Connected Component of each network.

To analyze the networks structure, we estimate the assortativity coefficient with respect to each attitude. The assortativity coefficient is the Pearson correlation coefficient of numerical attributes between pairs of linked nodes (this numerical attributes are the attitudes given by the model). It measures the similarity of connections in the graph with respect to the given numeric attribute [41]. Hence, the assortativity coefficient measures whether people relations are homophilic with respect to attitude. This behavior is commonly found in networks [5], and it has been documented in Twitter political discussion [10], including in Chile [20].

Figure 3: Most associated words to each attitude according to the TS-NMF model. Note that only single words are displayed, to avoid repetition in n-grams.

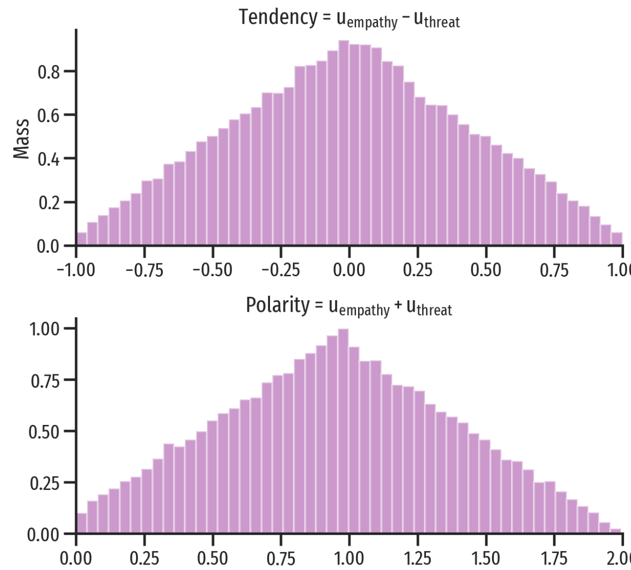


Figure 4: Top: tendency distribution for users. Bottom: polarity distribution for users.

In the next section we apply this methodology to the data set described in Section 4, covering an entire year of discussion about immigration in Chile.

6 RESULTS

Here we present the results of applying the methodology from Section 5 to the data set from Section 4.

Term Associations. Figure 3 shows the association of words with each attitude, empathy on top, threat on bottom. One can see that words associated to empathy include "integración" (integration), "salud" (health), and "educación" (education), reflecting their empathetic attitude. Words associated to threat include "delincuentes" (delinquents), "control" (control), and "ilegales" (illegals), reflecting a feeling of threat. Also, empathy group uses the word "Migrantes" (migrants) and threat group uses "Inmigrantes" (immigrants), which can be interpreted as that the empathy group is concerned about the general phenomenon (migration includes emigration and immigration), while the threat group only for the particular phenomenon (immigration).

Tendency and Polarity. Figure 4 shows the distribution of tendency and polarity for users. One can see that the distributions are fairly symmetric, with peaks in the center of the distribution. Figure 5 shows the tendency and polarity of tweets during the year under study, estimated using LOWESS. One can see that the tendency trend exhibits two interesting periods, before and after the news about the Leprosy case of an Haitian in July 31th. In the first period, tendency is slightly negative (threat), with an arguably low variability. In the second period, variability increases, and a small negative trend appears, even though at a point in time it reaches its maximum value (*i.e.*, maximum empathy) at the beginning of October. This could be explained by a news event reported in October 6th, about a Colombian citizen that gave birth on the street because a taxi driver expelled her from his car.

It is interesting that both news are related with the Integrated Threat theory and Intergroup Contact theory, respectively. On the one hand, the first event shows the immigrant as a threat, being a possible source of contagion of a disease (Leprosy). On the other hand, the second event shows the immigrant being a victim of violence and discrimination, which arguably makes people more empathetic. Regarding polarity, the trend exhibits a gradual increase in time, with two interesting peaks. The first one reflects the Leprosy case, and the second one reflects the presidential elections, where migration was a common topic in discussion.

LIWC Analysis. Figure 6 shows the differences of cognitive and emotional categories from LIWC in tweets grouped by tendency: *empathy* contains all tweets with tendency ≥ 0; *threat*, otherwise. For each category and group, we estimated the z-score for all tweets each month. As a general observation, one can see that both groups tend to have opposite behaviors. For instance, tweets in the empathy group are positively associated to the *sociability*, *family*, and *positive emotions* category more than tweets in the threat group. Conversely, tweets in the threat group are positively associated with *money*, *job*, and *inhibition* categories. This could be explained by the threat theory, as immigrants can be perceived as an economic threat and labor competition. Also, inhibition category can be interpreted by the desire to prohibit the arrival of more immigrants or to prevent them from accessing social benefits.

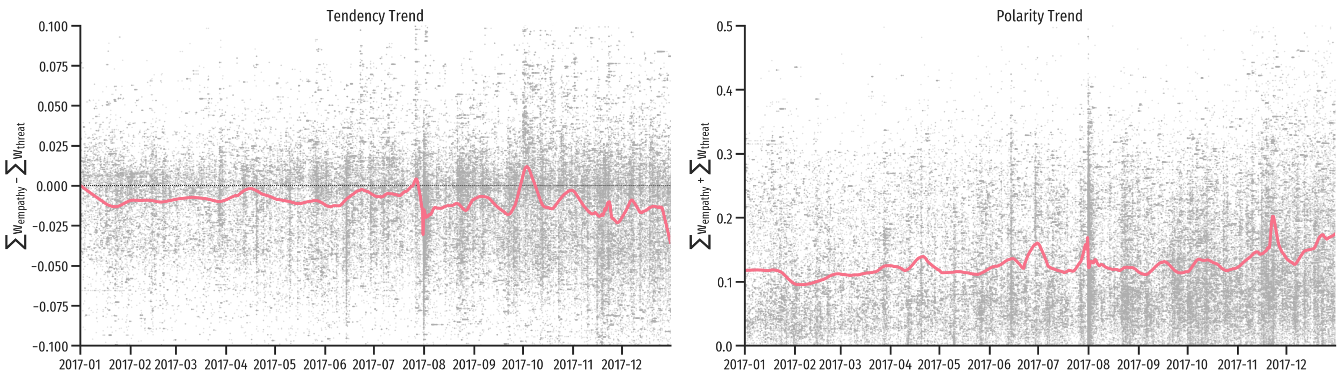


Figure 5: Trend distributions (top: tendency, bottom: polarity) for all tweets in the data set. Each tweet is a point, the x-position encodes its publication date, the y-position encodes its tendency or polarity. The line is the LOWESS interpolation of tendency and polarity.

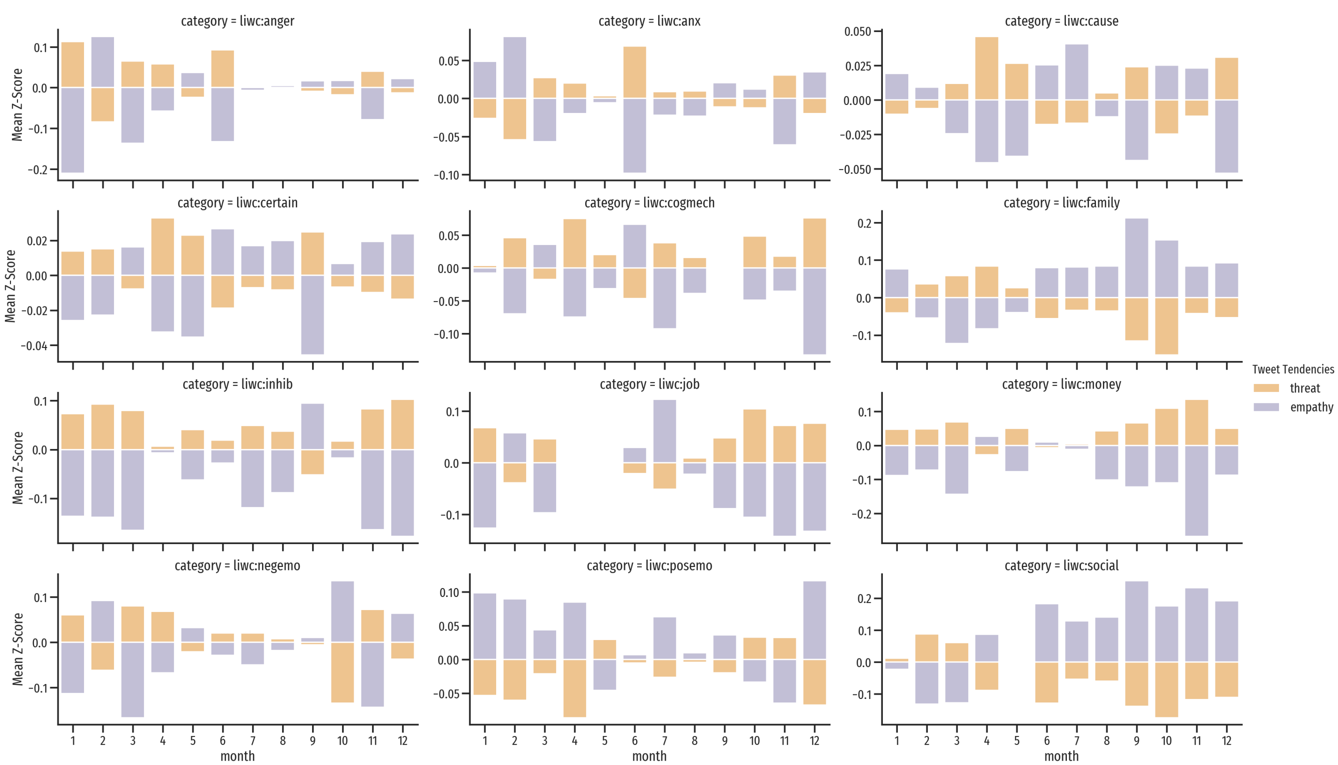


Figure 6: Association between attitudes (*empathy* and *threat*) and LIWC categories, per month. Each bar represents the association between groups, estimated with z-scores of fraction of words from each LIWC category and all other words. Purple bars indicate empathy associations, orange bars indicate threat associations.

Mention and Retweet Networks. The largest SCC of the retweet network has $1,239$ nodes and $6,441$ edges, while the largest SCC of the mention network has $1,868$ nodes and $10,201$ links. Figure 7 visualizes both networks using Hierarchical Edge Bundling [29]. This method allows us to make explicit the adjacency relations between users, as similar edges are bundled to decrease visual clutter. In the figure, each link is colored according to tendency of the source node (purple: empathy group, orange: threat group). Note that the visual encoding makes explicit the community structure in the retweet network and the heterogeneity of the mention network.

The assortativity coefficient for the retweet network are 0.26 (empathy) and 0.14 (threat), implying that homophilic behavior exists, but it is not as strong as in other topics (for instance, the discussion about abortion in Chile is greater [20]), and it is not

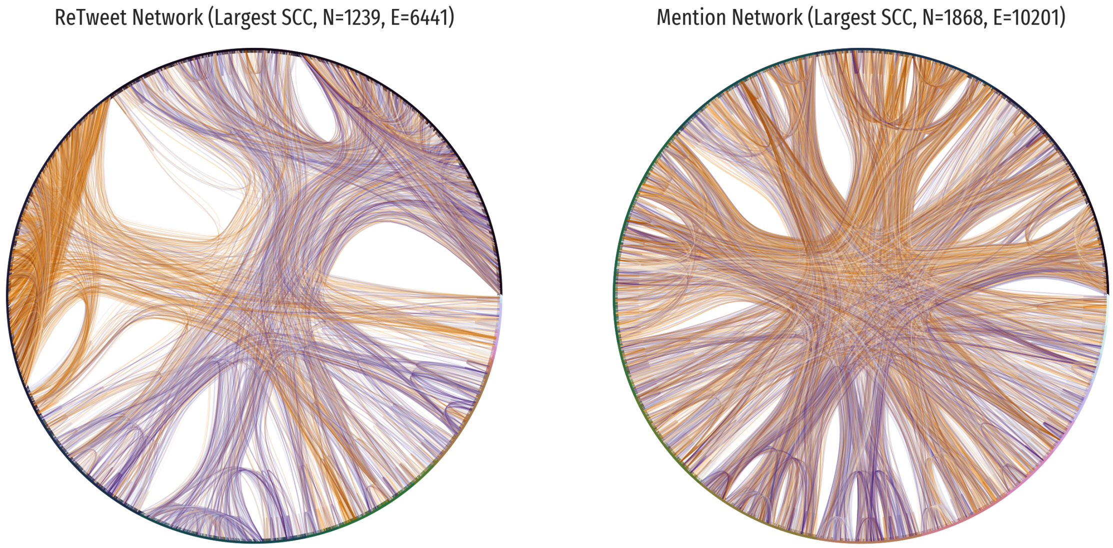


Figure 7: ReTweet Network (left) and Mention Network (right). Each node is a circle in the outside, sorted according to the connectivity patters to other nodes. Edges are lines that join nodes, where color is the attitude of the source node (purple: empathy, orange: threat). This encoding allows to group edges that are similar in terms of connectivity between groups.

equal in both groups. As hinted by the visualization, in the mention network the results are small: 0.06 (empathy) and 0.08 (threat). Thus, the retweet network is more segregated than the mention network. This could be explained because retweets are expected to be seen by all followers, and are a key factor in information diffusion, while mentions and replies are not. For instance, one user may send tweets to another holding an opposite position, but if there is no reply, then the interaction is not meaningful.

7 DISCUSSION AND FUTURE WORK

Migration is a controversial issue in Chile, and, although there are some studies about Chileans attitudes toward immigration [9, 37], they do not cover recent migration patterns. To complement knowledge about this topic, we defined a way to classify and measure attitudes, enabling to study the dynamics of perception with respect to immigration and performed a descriptive study of how immigration is perceived in Chile, according to Twitter discussion.

Our results may inform policy and intervention design, as it quantifies how people feel and communicate with respect to immigration. This is relevant, as there exists several contact strategies to improve relationships between social groups [44]. For instance, the discussion we analyzed is mostly targeted at Haitian migration. A majority of them is from Afro-Haitian descent, an ethnicity that was almost non-present in Chile.

There are two key aspects that need further exploration, and that limit the scope of our results: the representativity of Twitter, and the validation of the TS-NMF model. In terms of representativity, Twitter is a biased sample of the population [3]. As such, our results only cover this sample, even though it is not know to which degree nor to which sub-populations it represents. Having these biases into account will surely improve the interpretation of results. However, one aspect that needs to be considered is that Twitter is within the most popular applications in Chile [18], and that it reflects some cultural aspects, such as the country's centralization [19]. In terms of validation, the lack of ground truth or approximate measures of the problem stands in the way of effectively measuring the model accuracy, leaving us only with a qualitative evaluation.

Besides working on the limitations of our approach, there are two lines of future work that we devise. On the one hand, it would be relevant to understand the relationship between attitudes and actual presence of immigrants in a place. This would provide a way to measure real and imagined threat attitudes [33]. On the other hand, there is a potential influence of news events in attitudes. Given the rise of *fake news* and *post-truth* media, this would provide a way to measure the effect of such phenomena on how people feel with respect to a specific issue, migration in this case.

8 CONCLUSIONS

In this paper, we have characterized attitudes toward immigration by locals in Chile. We used a semi-supervised topic modeling technique (TS-NMF [39]) to identify attitudes grounded in two social theories, the Intergroup Contact Theory [1], and the Integrated Threat Theory [40, 49]. Then, we measured differences in attitudes using psycho-linguistic lexicons and interaction networks. As result, we found consistent behaviour with respect to social theory. There is still work to do in the evaluation and representativeness of our model, including the definition of a suitable ground-truth perception to validate our proposal. We believe our results help to inform the design of public policy and interventions to improve relations between groups in a country.

REFERENCES

[1] Gordon Willard Allport, Kenneth Clark, and Thomas Pettigrew. 1954. The nature of prejudice. (1954).
[2] Yair Amichai-Hamburger and Katelyn YA McKenna. 2006. The contact hypothesis reconsidered: Interacting via the Internet. *Journal of Computer-Mediated Communication* 11, 3 (2006), 825–843.
[3] Ricardo Baeza-Yates. 2018. Bias on the web. *Commun. ACM* 61, 6 (2018), 54–61.
[4] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. 2011. *Modern Information Retrieval: the concepts and technology behind search, 2nd. Edition.* Addison-Wesley, Pearson.
[5] Pablo Barberá. 2015. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis* 23, 1 (2015), 76–91.
[6] David M Blei. 2012. Probabilistic topic models. *Commun. ACM* 55, 4 (2012), 77–84.
[7] Kendrick T Brown, Tony N Brown, James S Jackson, Robert M Sellers, and Warde J Manuel. 2003. Teammates on and off the Field? Contact with black teammates and the racial attitudes of white student athletes 1. *Journal of applied social psychology* 33, 7 (2003), 1379–1403.
[8] Peter Burns and James G Gimpel. 2000. Economic insecurity, prejudicial stereotypes, and public opinion on immigration policy. *Political science quarterly* 115, 2 (2000), 201–225.
[9] Héctor Carvacho. 2010. Ideological configurations and prediction of attitudes toward immigrants in Chile and Germany. *International Journal of Conflict and Violence (IJCV)* 4, 2 (2010), 220–233.
[10] Michael Conover, Jacob Ratkiewicz, Matthew R Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. Political polarization on Twitter. *Proc. of ICWSM* 133 (2011), 89–96.
[11] Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in Twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality.* 51–60.
[12] Richard J Crisp and Rhiannon N Turner. 2009. Can imagined interactions produce positive perceptions?: Reducing prejudice through simulated social contact. *American psychologist* 64, 4 (2009), 231.
[13] Kareem Darwish, Walid Magdy, Afshin Rahimi, Timothy Baldwin, and Norah Abokhodair. 2017. Predicting Online Islamophopic Behavior after# ParisAttacks. *The Journal of Web Science* 3, 1 (2017).
[14] John F Dovidio, Anja Eller, and Miles Hewstone. 2011. Improving intergroup relations through direct, extended and other forms of indirect contact. *Group processes & intergroup relations* 14, 2 (2011), 147–160.
[15] Victoria M Esses, John F Dovidio, Lynne M Jackson, and Tamara L Armstrong. 2001. The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social issues* 57, 3 (2001), 389–412.
[16] Ruth Garcia-Gavilanes, Daniele Quercia, and Alejandro Jaimes. 2013. Cultural dimensions in Twitter: Time, individualism and power. *Proc. of ICWSM* 13 (2013).
[17] Roberto González-Ibánez, Smaranda Muresan, and Nina Wacholder. 2011. Identifying sarcasm in Twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2.* Association for Computational Linguistics, 581–586.
[18] Eduardo Graells-Garrido, Diego Caro, Omar Miranda, Rossano Schifanella, and Oscar F Peredo. 2018. The WWW (and an H) of Mobile Application Usage in the City: The What, Where, When, and How. In *Companion of the The Web Conference 2018 on The Web Conference 2018.* International World Wide Web Conferences Steering Committee, 1221–1229.
[19] Eduardo Graells-Garrido and Mounia Lalmas. 2014. Balancing diversity to counter-measure geographical centralization in microblogging platforms. In *Proceedings of the 25th ACM conference on Hypertext and social media.* ACM, 231–236.
[20] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. 2015. Finding intermediary topics between people of opposing views: a case study. In *Social Personalisation & Search*, Christoph Trattner, Denis Parra, Peter Brusilovsky, and Leandro Balby Marinho (Eds.). CEUR, Santiago, Chile.
[21] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. 2016. Encouraging diversity-and representation-awareness in geographically centralized content. In *Proceedings of the 21st International Conference on Intelligent User Interfaces.* ACM, 7–18.
[22] Shang E Ha. 2010. The consequences of multiracial contexts on public attitudes toward immigration. *Political Research Quarterly* 63, 1 (2010), 29–42.
[23] Jens Hainmueller and Michael J Hiscox. 2007. Educated preferences: Explaining attitudes toward immigration in Europe. *International organization* 61, 2 (2007), 399–442.
[24] Jens Hainmueller and Michael J Hiscox. 2010. Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American political science review* 104, 1 (2010), 61–84.
[25] Gordon H Hanson, Kenneth Scheve, and Matthew J Slaughter. 2007. Public finance and individual preferences over globalization strategies. *Economics & Politics* 19, 1 (2007), 1–33.
[26] GACCT Harman and Mark H Dredze. 2014. Measuring post traumatic stress disorder in Twitter. *In ICWSM* (2014).
[27] Gregory M Herek. 1986. The instrumentality of attitudes: Toward a neofunctional theory. *Journal of Social Issues* 42, 2 (1986), 99–114.
[28] Gregory M Herek and John P Capitanio. 1996. "Some of my best friends" Intergroup Contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality and social psychology bulletin* 22, 4 (1996), 412–424.
[29] Danny Holten. 2006. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on visualization and computer graphics* 12, 5 (2006), 741–748.
[30] Daniel J Hopkins. 2010. Politicized places: Explaining where and when immigrants provoke local opposition. *American political science review* 104, 1 (2010), 40–60.
[31] National Statistics Institute (INE). 2018. Síntesis resultados Censo 2017. urlhttps://www.censo2017.cl/descargas/home/sintesis-de-resultados-censo2017.pdf.
[32] Seth K Jolly and Gerald M DiGiusto. 2014. Xenophobia and immigrant contact: French public attitudes toward immigration. *The Social Science Journal* 51, 3 (2014), 464–473.
[33] Jeffrey S Kopstein and Jason Wittenberg. 2009. Does familiarity breed contempt? Inter-ethnic contact and support for illiberal parties. *The Journal of Politics* 71, 2 (2009), 414–428.
[34] Adam DI Kramer. 2010. An unobtrusive behavioral model of gross national happiness. In *Proceedings of the SIGCHI conference on human factors in computing systems.* ACM, 287–290.
[35] Onur Kucuktunc, B Barla Cambazoglu, Ingmar Weber, and Hakan Ferhatosmanoglu. 2012. A large-scale sentiment analysis for Yahoo! answers. In *Proceedings of the fifth ACM international conference on Web search and data mining.* ACM, 633–642.
[36] Fabio Lamanna, Maxime Lenormand, María Henar Salas-Olmedo, Gustavo Romanillos, Bruno Gonçalves, and José J Ramasco. 2018. Immigrant community integration in world cities. *PloS one* 13, 3 (2018), e0191612.
[37] Duncan Lawrence. 2015. Crossing the Cordillera: immigrant attributes and Chilean attitudes. *Latin American Research Review* 50, 4 (2015), 154–177.
[38] Daniel D Lee and H Sebastian Seung. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 6755 (1999), 788.
[39] Kelsey MacMillan and James D Wilson. 2017. Topic supervised non-negative matrix factorization. *arXiv preprint arXiv:1706.05084* (2017).
[40] Todd D Nelson. 2009. *Handbook of prejudice, stereotyping, and discrimination.* Psychology Press.
[41] Mark EJ Newman. 2003. Mixing patterns in networks. *Physical Review E* 67, 2 (2003), 026126.
[42] Josef Novotny and Filip Polonsky. 2011. The Level of Knowledge about Islam and Perception of Islam among Czech and Slovak University Students: does Ignorance Determine Subjective Attitudes? *Sociologia* 43, 6 (2011), 674–696.
[43] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71, 2001 (2001), 2001.
[44] Thomas F Pettigrew and Linda R Tropp. 2006. A meta-analytic test of intergroup contact theory. *Journal of personality and social psychology* 90, 5 (2006), 751.
[45] Daniele Quercia, Jonathan Ellis, Licia Capra, and Jon Crowcroft. 2012. Tracking gross community happiness from tweets. In *Proceedings of the ACM 2012 conference on computer supported cooperative work.* ACM, 965–968.
[46] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. 2011. Our Twitter profiles, our selves: Predicting personality with Twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.* IEEE, 180–185.
[47] Kenneth F Scheve and Matthew J Slaughter. 2001. Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics* 83, 1 (2001), 133–145.
[48] Paul M Sniderman, Pierangelo Peri, Rui JP de Figueiredo Jr, and Thomas L Piazza. 2000. The Outsider: Politics and Prejudice in Italy.
[49] Cookie White Stephan and Walter S Stephan. 2013. An integrated threat theory of prejudice. In *Reducing prejudice and discrimination.* Psychology Press, 33–56.
[50] Walter G Stephan and Krystina Finlay. 1999. The role of empathy in improving intergroup relations. *Journal of Social issues* 55, 4 (1999), 729–743.
[51] Walter G Stephan and Cookie White Stephan. 1985. Intergroup anxiety. *Journal of social issues* 41, 3 (1985), 157–175.
[52] Karolina Sylwester and Matthew Purver. 2015. Twitter language use reflects psychological differences between democrats and republicans. *PloS one* 10, 9 (2015), e0137422.
[53] Fiona A White and Hisham M Abu-Rayya. 2012. A dual identity-electronic contact (DIEC) experiment promoting short-and long-term intergroup harmony. *Journal of Experimental Social Psychology* 48, 3 (2012), 597–608.

applied sciences

MDPI

Article

A Framework to Understand Attitudes towards Immigration through Twitter

Yerka Freire-Vidal [1,*], Eduardo Graells-Garrido [1,2] and Francisco Rowe [3]

[1] Social Complexity Research Center, Universidad del Desarrollo, Las Condes, Santiago 7610658, Chile; egraells@udd.cl
[2] Barcelona Supercomputing Center (BSC), 08034 Barcelona, Spain
[3] Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool, Liverpool L69 7ZT, UK; f.rowe-gonzalez@liverpool.uk
* Correspondence: yfreirev@udd.cl

Abstract: Understanding public opinion towards immigrants is key to prevent acts of violence, discrimination and abuse. Traditional data sources, such as surveys, provide rich insights into the formation of such attitudes; yet, they are costly and offer limited temporal granularity, providing only a partial understanding of the dynamics of attitudes towards immigrants. Leveraging Twitter data and natural language processing, we propose a framework to measure attitudes towards immigration in online discussions. Grounded in theories of social psychology, the proposed framework enables the classification of users' into profile stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and temporal stance trends. We use a Twitter sample composed of 36 K users and 160 K tweets discussing the topic in 2017, when the immigrant population in the country recorded an increase by a factor of four from 2010. We found that the negative attitude group of users is smaller than the positive group, and that both attitudes have different distributions of the volume of content. Both types of attitudes show fluctuations over time that seem to be influenced by news events related to immigration. Accounts with negative attitudes use arguments of labour competition and stricter regulation of immigration. In contrast, accounts with positive attitudes reflect arguments in support of immigrants' human and civil rights. The framework and its application can inform policy makers about how people feel about immigration, with possible implications for policy communication and the design of interventions to improve negative attitudes.

Keywords: social network analysis; attitude classification; psycholinguistic analysis; public policy; migration

check for updates

Citation: Freire-Vidal, Y.; Graells-Garrido, E.; Rowe, F. A Framework to Understand Attitudes towards Immigration through Twitter. *Appl. Sci.* **2021**, *11*, 9689. https://doi.org/10.3390/app11209689

Academic Editor: Carlos A. Iglesias and Antonio Moreno

Received: 23 July 2021
Accepted: 17 September 2021
Published: 18 October 2021

1. Introduction

International migration has emerged as a major divisive global political and societal issue during the 21st Century, with increasing expressions of anti-migration sentiment [1]. Immigration has been portrayed as a major threat to social cohesion, notably during the UK Brexit Referendum and Trump presidential campaign, and drawn attention towards more restrictive migration policies, particularly in Western European countries and the United States [2–4]. Immigration sentiment is also an essential component for successful migrant integration into receiving societies. Discrimination, intolerance and xenophobia can hinder immigrants' capacity to secure employment, housing and achieve a sense of belonging in local communities, contributing to less cohesive societies [5–7]. Global initiatives have been established through the United Nations' Sustainable Development Goals (goal 10) [8] and Global Compact for Safe, Orderly and Regular Migration (goals 16 and 17) [9] to tackle anti-immigrant behaviour and thus facilitate migration integration.

The anti-migration sentiment is generally shaped by misconception [1], and social media has become a key channel to spread misinformation, contributing to the formation of misconceptions and manifestation of discriminatory acts against immigrants [10]. However, evidence from experimental study designs have revealed that attitudes can be shifted towards a more supportive view of immigration by explicitly addressing misconceptions [11]. Timely access and understanding of public opinion towards migration is thus critical for tackling misconceptions and understanding shifts in local openness to immigrants [12].

Empirical studies on attitudes towards immigration typically draw on survey questions about existing levels of immigration [10]. While surveys are useful, they are an expensive resource in terms of financial cost, labour and time. Moreover, considerable latency may impact data releases, impairing our ability to regularly monitor changes in migration sentiment, identify and tackle prejudice comments against immigrants. However, we know that anti-migration sentiment and prejudice comments can surge during economic recessions [13] and pandemics [14].

Digital trace data sources can now be used to complement and address some of the shortcomings of traditional survey data. Social media platforms, for example microblogging sites such as Twitter, offer a major source of real-time information to understand and quantify attitudes towards immigration. Twitter not only serves as a public forum to exchange opinions and ideas on a broad set of societal issues, including political events [15], abortion legislation [16,17] and migration [18–22], it also shapes the opinions of its users [23]. However, Twitter has also enabled the spread of misinformation and negative rhetoric fueling hate speech [24–26]. Such content has the potential to cause harm to individuals. It often translates into social tension outside the digital world and has played a role in the spread of hate speech during the COVID19 pandemic [14]. With this context in mind, we aim to answer the following research questions: (RQ1) Can we identify, quantify and classify attitudes towards immigration from social network data? (RQ2) What characteristics differentiate the content emitted by users with different attitudes? (RQ3) What emotions and psycholinguistic categories differentiate attitudes?

Using Twitter data and machine learning techniques, we aim to develop a replicable analytical framework to measure and analyse attitudes towards immigration. Specifically, we propose a framework to: (1) identify users' profile stances of positive and negative attitudes towards immigration; (2) analyse the content and psycholinguistic compositions of these profiles; and, (3) monitor their publication activity rhythm over time. We draw on a sample of 160 K tweets and 36 K users discussing immigration in Chile during 2017 when the immigrant population in the country was recorded to have increased by a factor of seven since 2002 from 105 k to 746 k, with over half of new arrivals occurring between 2012 and 2017 [27].

Our contributions are three-fold: first, we propose a methodological framework to operationalise mainstream theories of social psychology on the formation of attitudes towards immigration using Twitter data. It enables identifying users' positive and negative stances on immigration based on their public opinions and characterising the content and psycholinguistic features of these stances. Second, our proposed methodology reveals how digital traces can be used to complement and augment traditional data sources by enabling understanding of short-term changes in attitudes towards immigration and multidimensional views of immigration sentiment. Finally, our case study provides valuable insights into our limited knowledge of the patterns, experiences and challenges of recently arrived immigrants in Chile. Specifically, our work offers insights into the formation of attitudes towards immigration in Chile during a period of large migration influx, largely related to human displacement in Colombia and an exodus from Venezuela.

The paper is structured as follows. Section 2 discusses the theoretical and conceptual background related to our research work. Section 3 describes the dataset used for analysis. Section 4 describes the proposed methodology before Section 5 presents the results. Section 6 discusses the implications and limitations of our work, and Section 7 offers some concluding remarks.

2. Background

In this section, we review theoretical approaches to the formation of attitudes towards immigration. Then, we describe the literature on measuring attitudes based on Twitter data to illustrate the key challenges and advantages of using this data source. Finally, we describe immigration in Chile.

2.1. Theories and Measurement of Attitude Formation

Two theoretical models are often used to describe the formation of attitudes towards immigration. These are the Intergroup Contact Theory (ICT) [28] and the Integrated Threat Theory (ITT) [29,30]. ICT explains how *positive* attitudes form, while ITT describes how *negative* attitudes are created.

The ICT postulates that increased social interaction among people from different population subgroups reduces prejudice and enhances trust. The theory is based on conditions such as common goals for different groups and the benefits of cooperation and support (both formal and informal) to reach those goals. Existing research has contributed evidence to support these arguments [31–33]. Increased intergroup interaction reduces fear and anxiety that may exist when people interact with individuals from an unfamiliar group [34,35] by promoting empathy and understanding [36,37].

In contrast, the ITT predicts that social interaction among people from different groups may lead to perceptions of threat and contempt toward members of the different groups [29,30]. Two types of threat describe the formation of negative attitudes: symbolic and realistic. In the context of migration, these threats are related to competition in the labour market, to public health concerns from possible diseases, to increased crime and physical well-being, and to perceptions of the size of the foreign group, among others [38,39], as well as an increased fiscal burden [40].

The common method to measure attitudes toward immigration is to use public opinion survey data. Some examples include the Gallup World Poll, the Pew Global Attitudes Survey, the International Social Survey Programme, the World Values Survey, the Iposos Global Trends, the European Social Survey, and the Eurobarometer [41]. Data from the Gallup World Poll revealed that in all major regions of the world, people are more likely to want immigration to remain at the current level or increase, rather than decrease, with the exception of Europe. However, there is great variability between countries. In Europe, for example, southern Europeans tend to show more negative attitudes towards immigration than northern Europeans, who show more positive attitudes [4].

Although there is now a large repository of public opinion surveys on migration, some important limitations with these traditional methods arise. For example, existing surveys are mainly from European and North American countries, the number of questions collecting opinions on migration is highly variable across countries, their frequency is coarse, and they are often costly to implement. Thus, it has become a necessity to obtain frequent, comprehensive information covering different regions of the world to understand attitudes toward immigration, as stated in the first objective of the Global Compact for Safe, Orderly and Regular Migration: "to collect and use accurate, disaggregated data to formulate evidence-based policies" [9].

In this work, we propose that the discussion in social networks offers a novel and inexpensive source to complement traditional data systems. It can be collected in near real-time, which would allow for richer analyses grounded on the aforementioned theories. In the following, we describe the potential of the social network Twitter in this regard.

2.2. Social Media Analysis in the Study of Human Behaviour

To discuss social media analysis from three perspectives: human behaviour, attitude classification and psycholinguistic analysis. We focus on the analysis of the micro-blogging platform Twitter.

Researchers from different disciplines have used data extracted from the Twitter platform to study psychological and socio-cultural characteristics of communities. Some examples include: personality differences between Democrats and Republicans [42], influence of diurnal and seasonal variability on mood [43], happiness associated with Christianity/atheism [44], political polarisation [45], influence of culture on personal actions [46], prediction of attitudes in response to a trigger event [47], measurement of the level of integration of immigrants in different cities of the world [48], among others. When the analysis is about groups of people, Twitter data highlights its potential use in the detection of hate speech [49], i.e., threatening, harassing or seriously offensive language, as well as to characterise hateful users [50]. The detection of hate speech against immigrants is a particular case of the general framework of hate speech detection. Several studies have used Twitter data to study hate speech against immigrants [18–22]. C. Arcila C. et al. [22] model and characterise anti-immigrant hate speech on Twitter in Spain. They find that hate speech against immigrants includes Islamophobia, rejection of public support towards the immigrant group, and has a greater presence of offensive than violent language.

In this context, we develop a mixed approach grounded on the theories about attitude formation instead of following a hate speech approach, which is limited to hate but not necessarily opposition/approval or feelings of threat/empathy toward migration. Each formation theory defines an attitude, and, in cases where the classifier confidence is low, we define an *undisclosed* stance to account for participation in the debate without disclosing attitude [51]. Particularly, we build upon our previous work to classify users into attitudes as political stances using a tree-based classifier [17]. As stances, attitudes are not always explicit, and thus, they must be predicted. Two types of features are commonly used for prediction. On the one hand, stance can be predicted using network interactions, based on the assumption that like-minded individuals are more likely to interact [52,53]. On the other hand, lexical analyses have shown to allow predicting stance as vocabularies within stances tend to have strongly associated words [54,55], and even other non-textual cues such as emojis [17].

A related area to attitude classification is sentiment analysis. In general, such methods extract the polarity of a text in terms of positive, neutral, and negative intents [56,57]. These measures may be correlated to attitudes, however, empathetic tweets about migration could speak negatively about discrimination, and threat-related tweets could speak positively about the country, making the interpretation of the results difficult. However, we perform one type of emotional analysis when characterising the content published by users in each attitude. We use the Linguistic Inquiry and Word Count (LIWC) psycholinguistic lexicon [58] to distinguish multidimensional aspects of the migratory discussion. These aspects include multiple emotions and other topical dimensions of expression, which may be compared between attitudes. LIWC usage with social media data includes identifying health issues [59,60], predicting political sentiment [61], and understanding perception about the transportation experience [62]. The Spanish version has been validated experimentally [63], enabling its consideration in our work.

2.3. Immigration in Chile

In the Latin American context, Chile presented the greatest increase in the weight of the immigrant population between the last two censuses [64]. Although immigration is not a new phenomenon in Chile, it has increased dramatically in recent years and has become a relevant issue in the national debate (see Figure 1). The immigrant population grew from 1.27% of the total population in 2002 to 4.35% according to the 2017 Census, with a more diverse origin for all immigrants than in previous years. As of 2017, 66.7% of all resident immigrants arrived in Chile from 2010, and 61% arrived in 2015–2017 (considering up to 19 April 2017, the date of application of the census).

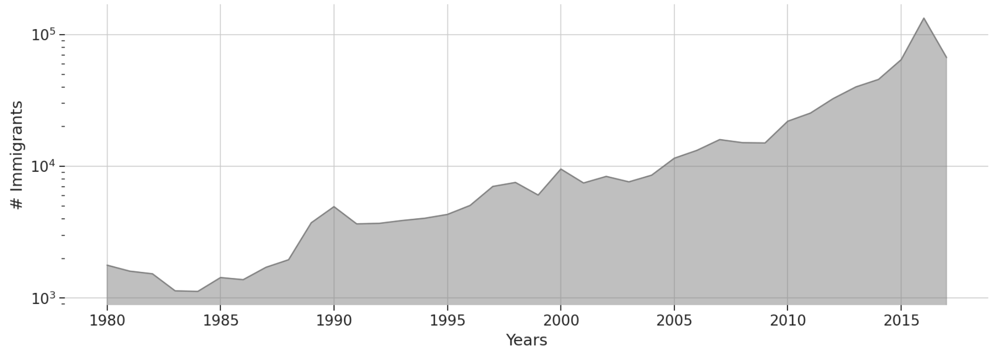


Figure 1. Number of immigrants in Chile by year of arrival according to the 2017 Census.

There were two milestones that impacted the national agenda and installed migration as a social concern [65]. First, in August 2016, the Chilean Public Ministry opened an investigation into the possible involvement of an airline in a case of migrant trafficking. Second, in July 2017, the Ministry of Health confirmed a case of leprosy in a Haitian citizen. Leprosy is a disease without records in the country except for overseas islands [66].

A national-level survey in Chile revealed that 57% of nationals agree that the country should take more drastic measures to exclude illegal immigrants, versus 19% who disagree. In addition, 41% of respondents agree that immigrants raise crime rates, compared to 38% who disagree with this statement. In the area of labour competition, 40% of respondents agree that immigrants take jobs away from Chilean-born people, and 36% disagree. This survey was applied in the months of April-May 2017 [67]. Arguably, the questions are more related to the ITT than the ICT. To the extent of our knowledge, there are no other surveys that try to understand how Chileans feel about migration. Hence, this scenario puts Chile as a relevant case study with respect to measuring local attitudes towards immigration. Such analyses may provide complementary knowledge of the attitudes and perceptions of the population, first, by including social theories, and second, by providing fine-grained, dynamic insights.

3. Dataset

In this section, we describe the Twitter platform and the dataset used to measure attitudes towards immigration.

Twitter is a micro-blogging platform where users report a screen name, a full name (which can be real or fictitious), a location (real, fictitious, or empty), and a small autobiography, among other attributes. Each user publishes posts (called *tweets*) with a maximum of 280 characters. Twitter also allows interaction between its users: users can *follow* others users and to see their tweets in their own timelines. Users can mention other users in their own tweets using a handle (e.g., @username), *quote* other tweets or adding commentary to them, or *retweet* (RT) another tweet to share it with one's audience.

To crawl tweets that discuss immigration in Chile, we connected to the Twitter Streaming API using a system designed to crawl Chilean tweets [68]. The query parameters were keywords related to immigration, e.g., "inmigración" (*immigration*), "inmigrante" (*immigrant*), "fronteras" (*borders*), "racismo" (*racism*), etc.; and origin countries with their respective denomyms. Thus, the dataset after data cleaning is composed by 160,775 tweets (54,252 are plain tweets and 106,532 are retweets—RTs–. In addition, 20,248 tweets are quotes and 19,265 are replies) that are on topic during 2017, written by 36,698 users (see Figure 2 for the temporal distribution of content).

The cleaning process ensured that the discussion under analysis was about human migration in Chile. Examples of noise topics were: racism toward indigenous groups in Chile, bird migration, a South American soccer championship, national presidential elections, migration issues in México, the U.S.A., and Spain, among others. Hence, we excluded tweets from users with a reported (or predicted) location different to Chile, as well as tweets in languages other than Spanish.

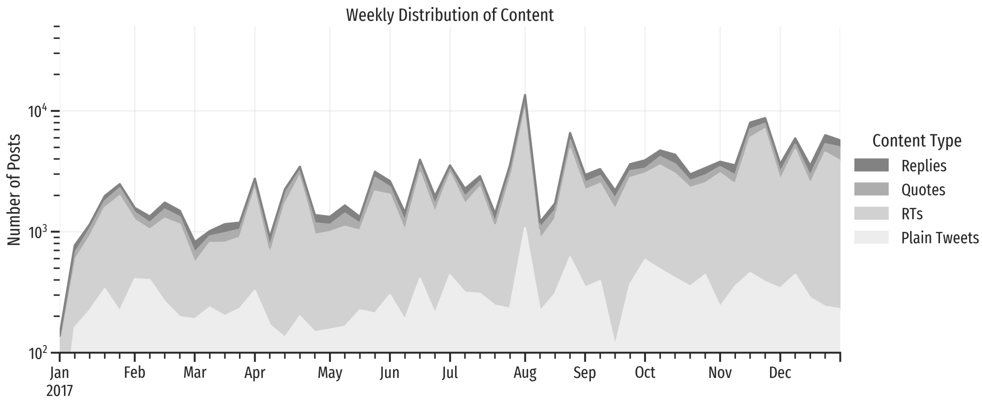


Figure 2. Weekly volume of content in the dataset, disaggregated by type of content/interaction.

4. Methodology

In this section, we describe how to classify, quantify and characterise the attitudes toward immigration. The methodology is composed of the following steps: theory-informed profile tagging, which enables to pinpoint some users with an attitude; propagation of user attitudes to the rest of the dataset, after training a classifier with the tagged users; and then perform a characterisation of attitudes from the lens of users, content, psycholinguistics, and dynamics. A schematic diagram of the methodology is shown in Figure 3.

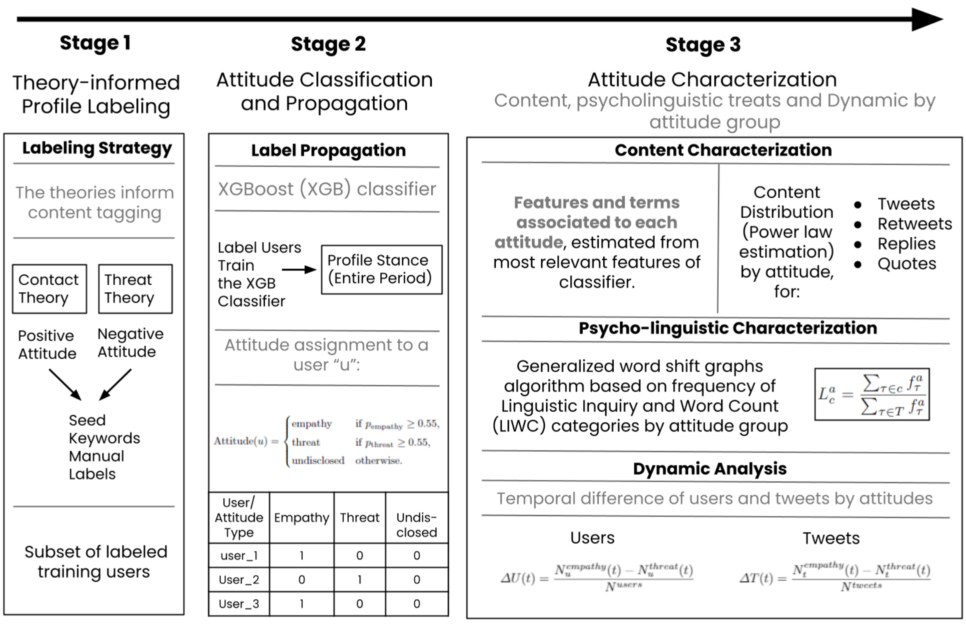


Figure 3. Schematic diagram of the proposed methodology.

4.1. Theory-Informed Profile Labelling

Here we define the two attitudes we classify profiles into, and how to label profiles according to these attitudes for classification. In the spectrum of attitudes towards immigration there are two that define two opposite extremes. Usually, these types of categories are named *positive/negative*, *in-favour/against*, or similar. Here we rely on the social psychology theories described in Section 2: the Intergroup Contact Theory (ICT), and the Integrated Threat Theory (ITT). Based on those theories, we name the two attitudes we are classifying as *empathy*, due to the empathy toward immigrants, and *threat*, due to perceptions of threats regarding migration.

Given the potential size of the discussion under analysis, manually labelling the user profiles (or their tweets) into these two categories is expensive and impractical. In view of this difficulty, we predict attitudes using a classifier trained on a labelled subset of the dataset. This subset is labelled automatically from a list of seed patterns and keywords for each attitude, as words are an effective mechanism to predict the community a user belongs to [18,69,70]. To identify seeds, we iteratively explore the dataset to seek for features that could be mapped to the *empathy* and *threat* attitudes. In *empathy*, we look for features that indicate that immigrants are welcome and will be received in equal conditions (e.g., "we are all immigrants"). In *threat*, we look for features that show that immigrants are not welcome or qualify them or their arrival negatively (e.g., "illegal immigrants take our jobs"). The labelled terms are not necessarily frequent, instead, they are discriminative, i.e., it is likely that someone in its corresponding category would use the term, and not from the other. The list is built iteratively in the sense of running the first steps from this section up to the classification step, and then exploring the usage of discriminative terms by people in each group to look for other potential seeds. See example seeds for each attitude in Table 1.

Table 1. Seed patterns for each attitude toward immigration.

Attitude	Seed Words and Hashtags (and Their Frequencies in Dataset)
Empathy	#bienvenidosachile (*welcome to Chile*, 26), #chilesinbarreras (*Chile without barriers*, 7), #chilediverso (*Diverse Chile*, 43), #noalaxenofobia (*no to xenophobia*, 126), #nomasxenofobia (*no more xenophobia*, 30), #racismo (*racism*, 1336), #stopracismo (*stop racism*, 48), #nomasracismo(*no more racism*, 10), #noalracismo (*no to racism* 73), #pongamonosinmigrantes (*let us become immigrants*, 67), #todossomoshermanos (*we are all siblings*, 157), #todossomosmigrantes (*we are all migrants*, 1108), #bienvenidosmigrantes (*welcome migrants*, 6), #ningunserhumanoesilegal (*no human being is illegal*, 16), #nadieesilegal (*no one is illegal*, 29), #nohayserhumanoilegal (*no human being is illegal*, 22), #redmigrante (*migrant network*, 105), #interculturalidad (*interculturality*, 218), #díadelmigrante (*migrant's day*, 269), #sinfronteras (*without borders*, 371), inhumano (*inhuman*, 349), multicultural (372), diversidad (*diversity*, 1531)
Threat	#inmigrantesilegales (*illegal immigrants*, 7), #nomasinmigrantes (*no more immigrants*, 41), #vendepatria (*sells homeland*, 5), #estadodecatastrofe (*state of catastrophe*, 44), invasión (*invasion*, 1258), invaden (*they invade*, 62), turba (*group of people generating chaos*, 732), prestamistas (*moneylenders*, 372), narcotráfico (*drug trafficking*, 531), turistas (*tourists*, 2186), fronterizo (*at the border*, 292), enfermedades (*diseases*, 2138), narcotraficantes (*drug dealers*, 291), expulsarlos (*eject them*, 105), echarlos (*to take them out*, 1172), deportarlos (*deport them*, 88)

Next, we label users who match these patterns. Those who are labelled in both attitudes have their labels removed. Note that we assume that in the period under study attitudes do not change. Additionally, we manually label accounts of institutions (such as the International Migration Office and the Jesuit Service for Migrants), opinion leaders, journalists, and politicians that have explicitly expressed their attitude on the issue.

4.2. Attitude Classification and Propagation

To predict attitudes, we follow a bootstrapped approach, where we propagate the user labels from the previous step to the rest of the dataset. We use the XGBoost classifier that trains decision trees using gradient boosting [71]. The input feature matrix is the concatenation of several matrices:

- A content-term matrix, where each row represents user i, and each term j can represent a word, hashtag, username, URL or emoji. Thus, a cell (i,j) contains the number of times user i has used the term j in their tweets.
- A profile-term matrix, analogous to the previous one, but this time for the terms contained in the full name and biographical self-description of each user.
- A profile-domain matrix, mapping to each user's home page its main domain (e.g., twitter.com) and their main top level domain (e.g., .com).
- Since homophily may vary or be absent in different interaction layers [72], we consider the three types of interaction separately. Thus, we build three adjacency matrices based on the interactions in the discussion: retweets, replies, and quotes. Each matrix stores in a cell (i,j), the number of times user i has interacted with user j (for instance, if i retweets j one time, $c(i,j) = 1$).

- A user–attitude interaction matrix for each type of interaction, where each cell contains the number of times the corresponding user has interacted with other users that were labelled with an attitude.

Then, we train the classifier using the set of labelled users. To avoid overfitting, we take two measures. First, the gradient boosting is performed with early stopping, using a validation set of 15% of the training observations. Second, we removed columns from the feature matrix that were used for labelling. This includes the seed keywords for each attitude, as they perfectly separate users from both groups and our goal is to classify users who do not use these terms in their content.

After having trained the classifier, we predict the attitude of the rest of the dataset. For a given profile u, the classifier outputs a value $p_a(u)$ for each attitude a that lies in $[0,1]$, corresponding to the fraction of decision trees that vote for the corresponding attitude. Note that the value of $p_a(u)$ is not a real probability. Thus, we apply a small threshold to consider predictions with a number of voters higher than a random choice. Those users who cannot be classified are marked as *undisclosed*. As a result of this stage, we assign an attitude to a user u according to the following function:

$$\text{Attitude}(u) = \begin{cases} \text{empathy} & \text{if } p_{\text{empathy}} \geq 0.55, \\ \text{threat} & \text{if } p_{\text{threat}} \geq 0.55, \\ \text{undisclosed} & \text{otherwise.} \end{cases}$$

After predicting attitudes, we manually check for profiles that are highly active/followed in the discussion and could have been mislabelled by the classifier. We add those manual labels and then repeat this stage.

4.3. Attitude Characterisation

In the last stage, we characterise each attitude from multiple perspectives. We describe what characterises each attitude from the lens of content (what is published in each category and how?), the lens of psycholinguistics (which semantic categories discriminate the expression of emotions in each category?), and the lens of dynamics (when does each attitude express their opinions?).

4.3.1. Content Characterisation

To measure what is published in each category and how, we focus on two aspects of content: the profile features that are most associated with each attitude, and the distribution of tweet vocabulary per attitude. On the one hand, the association of the features deemed important for classification provides insight into which ones are associated with each attitude [17]. However, the classifier estimates plain feature relevance, without any association to each attitude. Hence, for all relevant features, we estimate their association to each attitude using the log-odds ratio with uninformative Dirichlet prior [73], a measure that weights features in a similar way to TF-IDF, with the addition of controlling the variability of frequency. We apply this weighting to three documents, one per attitude (including *undisclosed*). Each document is the column-oriented sum of the feature matrix for all users predicted in the corresponding attitude. Here, we expect to find different terms that express the same concept but are associated with opposite attitudes [70], as well as differences in features from the self-reported biographies, and even in emoji usage [17].

On the other hand, we know that the content generation in Twitter tends to follow a powerlaw [74,75], that is, their distributions can be described approximately as $P(x) \sim x^{-\alpha}$. Hence, for each attitude, we characterise the volume of content (the number of plain tweets, RTs, quotes, and replies) according to the exponent α of their fitted powerlaw distributions. These exponents enable us to compare if attitudes behave differently in their discussion and interaction mechanisms.

4.3.2. Psycholinguistic Characterisation

The previous lens provided differences in the content published by each attitude. However, a direct content-based approach does not discriminate the expression of emotions in each category. To gain a deeper understanding of attitudes in this direction, we use a well-known psycholinguistic lexicon named "Linguistic Inquiry and Word Count" (LIWC) [76]. LIWC was designed to capture emotional, cognitive, and structural components present in text. It is available in several languages. Its Spanish version contains 7515 words classified in one or more of 72 categories that belong in four dimensions:

1. Standard linguistic processes: articles, prepositions, pronouns, etc.
2. Psychological and affective processes: positive and negative emotions, with subcategories such as anger and anxiety.
3. Relativity: time, verb tense, motion, space.
4. Personal matters: sex, death, home, occupation, etc.

We will focus our analysis on LIWC categories that are possibly associated with some of the factors shaping empathetic and threatening attitudes towards immigration. With regard to *empathy* attitudes, we consider the following LIWC categories as relevant: affective processes, positive feeling and emotions, optimism and energy, humans, social, family, and inclusion. With regard to *threat* attitudes, we consider the following LIWC categories as relevant: anger, anxiety, negative emotions, inhibition, death, body, job, and money. We estimate the association between attitudes and LIWC categories. First, we define the association of LIWC category c to attitude a, L_c^a, as the relative frequency of words in c with respect to all terms in the discussion per attitude:

$$L_c^a = \frac{\sum_{\tau \in c} f_\tau^a}{\sum_{\tau \in T} f_\tau^a},$$

where T denotes the vocabulary, and f_τ^a the total frequency of vocabulary term τ by accounts with attitude a. To explore these associations, we visualise them using Generalised Word Shift Graphs [77]. These visualisations summarise each attitude according to the differences in association between attitudes and LIWC categories, with the aim of describing the emotional and semantic aspects of attitude formation.

4.3.3. Dynamics and Events

Our final lens of characterisation aims to understand when each attitude expresses its opinions. Although we assume that attitudes are constant in the time under study, there still could be differences regarding how they are expressed in time in terms of volume. Particularly, we compute the daily difference between the number of tweets and the number of users per attitude. Positive values of this difference indicate a tendency towards *empathy*, whereas negative values indicate a tendency towards *threat*. Then, we compare these time-series with a null-model where we randomly permute the predicted attitudes for users 1 K times. If for a given day, the time-series does not intersect the 95% confidence interval of the null series, we consider it different with statistical significance. Finally, by identifying dates with salient and significant differences, we can explore what triggers the expression of each attitude. In this work, we manually do this through visual exploration.

The proposed methodology provides a full pipeline to characterise users attitudes towards immigration in a micro-blogging platform. Next, we apply this methodology to a set of tweets in Chile during 2017.

5. Results

This section first presents the results associated with the classification and characterisation of user attitudes towards immigration according to *empathetic* attitudes and *threatening* attitudes. Next, we analyse the content composition of these user profiles before examining differences in their lexical expressions and temporal fluctuations in attitude sentiment.

5.1. Attitude Identification and Classification

To answer our question RQ1, we performed an exhaustive search in the dataset for hashtags and words that could be strongly associated with each attitude (see Table 1 for more examples). For instance, the hashtags #bienvenidosachile (welcome to Chile) and #inmigrantesilegales (illegal immigrants) can be directly mapped to each attitude. We found more hashtags in the *empathy* group than in the *threat* group. When tagging threat users, we included words that are not normally associated with threats or negative comments against immigrants. For instance, "turistas" (*tourists*) is not a word commonly associated with a threat attitude. However, in our dataset, tourists tended to be used to express anger about people entering the country on a tourist visa and overstaying.

In total, 3.1 K accounts were tagged in the empathy group, and 1.2 K accounts were tagged in the threat group (see Table 2). The attitude classifier presented good performance, with high precision for both attitudes (0.95 in *empathy*, 0.81 in *threat*) and high recall (0.88 in both). Such good performance is expected, due to a tagging strategy based on a perfect separation of accounts from each attitude. However, this strategy works well identifying a tendency towards attitudes because the keywords used for tagging are excluded from the learning process [17].

Table 2. Performance (precision and recall) of the user profile classifier based on 10-fold cross-validation.

Attitude	Precision (mean)	Precision (std.)	Recall (mean)	Recall (std.)	Labelled Accounts
Empathy	0.95	0.04	0.88	0.13	3118
Threat	0.81	0.16	0.88	0.12	1233

Figure 4 presents the top-50 features according to their importance in classification. These features include: retweets to accounts that were be automatically tagged with each attitude, mentions to then-president Michelle Bachelet (@mbachelet), attitude-relevant terms such as cesantía (*unemployment*), delincuentes (*criminals*), inmigración (*immigration*), migrantes (*migrants*), and haitianos (*Haitians*); and mentions to migrant-support institutions as @sjmchile (Jesuit Center for Migration). The remaining features in the top-50 include interactions with authors of viral tweets, and terms relating to specific content, such as a quote by Umberto Eco.

We then used the classifier to propagate attitudes to the rest of the dataset (88% of the total dataset). Table 3 reports the resulting classification. Considering the labelling and propagation processes, a total of 71.98% user profiles were classified in *empathy*, 3.24% in *undisclosed*, and 24.78% in *threat*. This is an unexpected outcome. We expected the distribution of user profiles to be negatively biased given the social divisive nature of immigration in Chile. However, this may imply that negative users are more vocal than positive users and hence there is a perception for negative comments to dominate the public discussion around immigration issues in Chile. Table 3 reveals this is the case with the threat group generating considerably more content per user (with a ratio of 3.5 tweets per user, for the empathy attitude group, versus 6.7 tweets per user, for the threat attitude group) than the empathy group, particularly quotes and replies.

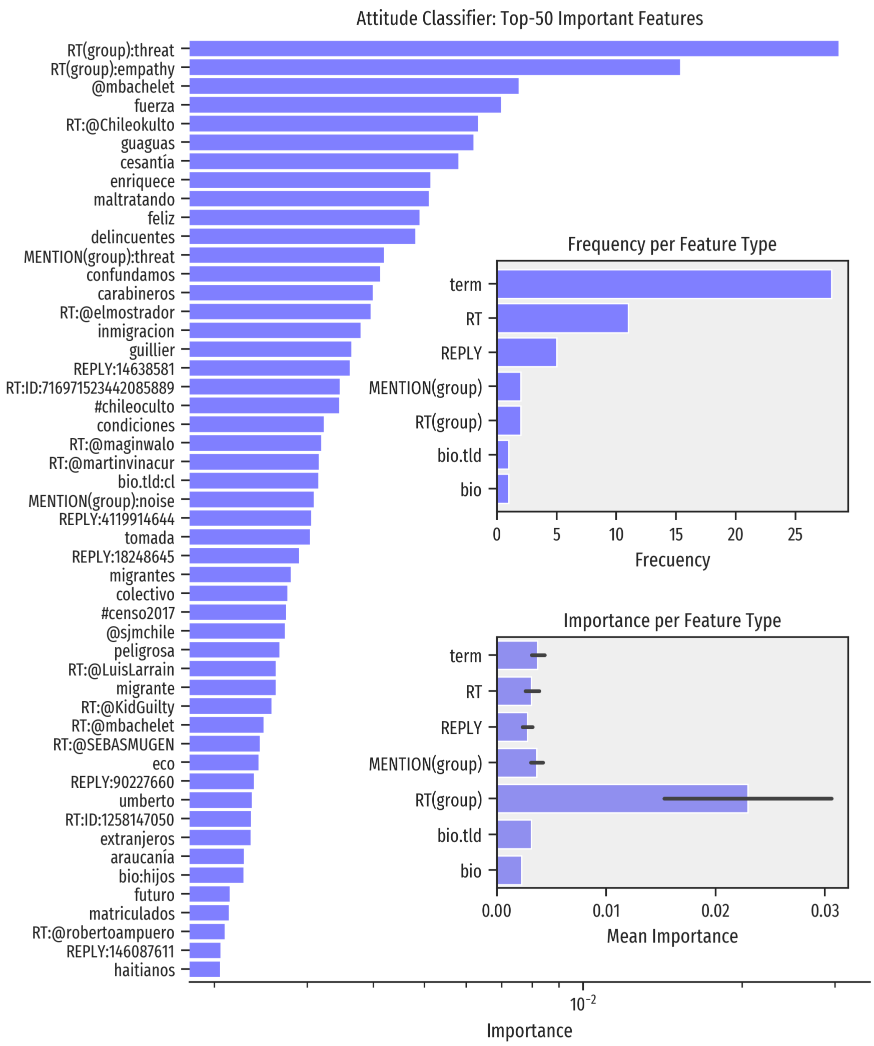


Figure 4. Attitude classification: Top 50 most important features.

Table 3. Distribution of user accounts and tweets per attitude and account type.

Attitude	Accounts	Total Tweets	RTs	Quotes	Replies
Empathy	% 71.98 (26,414)	% 57.46 (92,374)	% 61.94 (65,989)	% 45.62 (9238)	% 32.81 (6320)
Undisclosed	% 3.24 (1190)	% 4.52 (7261)	% 1.44 (1531)	% 1.87 (379)	% 3.61 (695)
Threat	% 24.78 (9094)	% 38.02 (61,140)	% 36.62 (39,012)	% 52.50 (10,631)	% 63.59 (12,250)

5.2. Content Characterisation

To better understand the discussion driving the expression of positive and negative attitudes towards immigration and answer our question RQ2, we performed content analysis producing a characterisation of the most frequently used terms associated with each user stance profile. We weighted each term using log-odds ratios with an Uninformative Dirichlet Prior [73]. Figure 5 displays the top-30 attributes per attitude, revealing that own-attitude retweeting behaviour is the most prominent feature for *empathy* and *threat*. This is arguably expected due to homophilic interactions [72], a phenomenon that has been observed in political discussion in Chile [78]. However, care must be taken when interpreting the result, as this reflects interactions with accounts that were pre-labelled only. A plausible explanation is that highly popular accounts that can be labelled through methods tend to make their positions explicit.

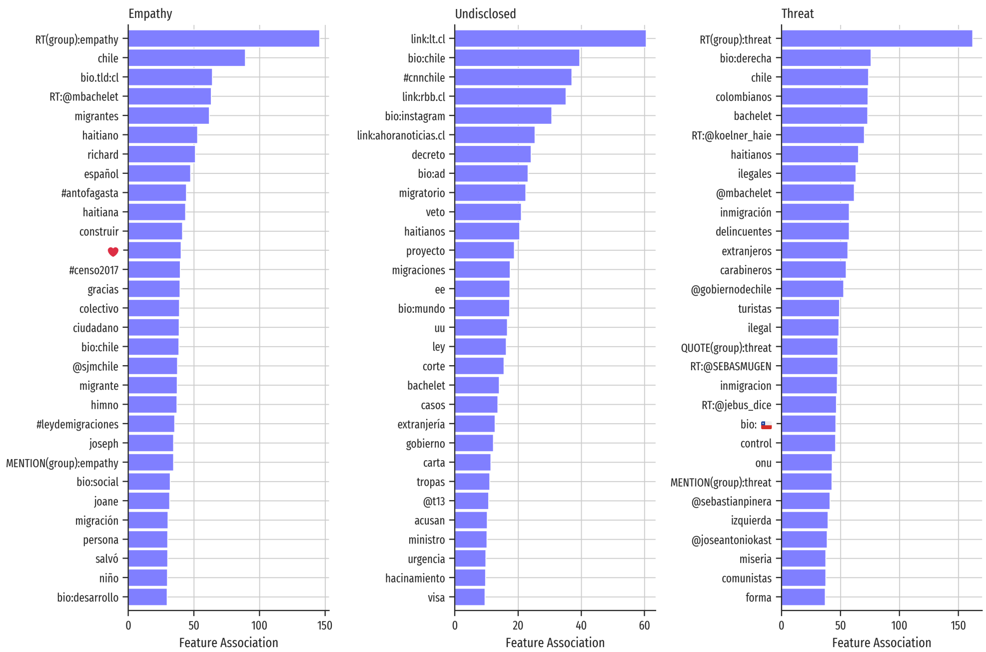


Figure 5. Top-term associations to each attitude (or lack thereof), estimated using log-odds ratio with uninformative Dirichlet prior.

Neutral terms, such as migrante (*migrant*) and migración (*migration*) and haitiano (*Haitian* as a singular noun) feature prominently in the *empathy* group. Emojis are typically used to express political attitudes [17] and we observe that the red heart emoji is associated with *empathy*, possibly as an expression of solidarity or liking of immigrants-related content. Content related to news items also feature prominently in empathetic Twitter content. Ranking in the top-50 items, we identified the case of Richard Joseph: a Haitian who saved a person that fell from a nine-story building; and, Joane Florvil: a Haitian mother who died after being arrested by the Chilean police under unclear circumstances. We also identified political-based debates, including retweets of Michelle Bachelet's and mentions of a debate of a national migration law (#leydemigraciones).

In contrast, words such as inmigración (*immigration*), extranjeros (*foreigners*) and haitianos (plural, *Haitians*) are often used by the threat group. Immigration placed the focus on migration as a process and concerns about its potential implications for the national health, education and labour market systems, rather than on understanding the individuals themselves. The use of foreigners and Haitians may be used to draw a clear distinction between the "we" and "us". The use of the Chilean flag also features prominently among the *threat* group, arguably associated with expressions of nationalism. *Threat* users also tend to identify their alignment with right-wing views, including "derecha" (right-wing) in their biographies; tagging right-wing politicians, such as Sebastián Piñera (@sebastianpinera, current president of Chile) and José Antonio Kast (@joseantoniokast, extreme right-wing presidential candidate); and, posts about left-wing parties ("ìzquierda") and the Communist Party ("comunistas"). Overall, the results highlight a strong association between anti-migration views and conservative political ideologies. This is despite not including any explicit political keywords in the seed list for our stance classifier.

We did not observe incidence of the seed words (used to train the classifier) in the content analysis and relevant features associated with each attitude; since only one of them appears within the top ranking (tourists, in the threat group). Note that the total number of training seed words is small (56 in total; 34 for the empathy attitude, and 22 for threat) and their frequency of use is mostly small (see Table 1 for some examples).

To identify and quantify differences in the diffusion of content generation, by both groups of attitudes, we analysed the respective distributions of the number of tweets, retweets, quotes and replies. Because other studies suggest that these measures of content generation follow a powerlaw distribution [75,79], we fit powerlaw regressions on the distributions of the aforementioned measures. Figure 6 reports the results. Theoretically, a symmetrically diagonal line would indicate perfectly equal spread generation of content. A perfectly horizontal line would indicate high concentration of the content generation by a single user. Figure 6 reveals a relatively symmetrical distribution in the generation of content for empathetic users across tweets, retweets, quotes and replies. This contrasts with the distributions associated with the threat group that displays higher concentration of content generation in a small number of users, with consistently lower powerlaw exponents to empathetic users. This is particularly prominent for retweets suggesting a high degree of interaction within the social network of threat users. These findings confirm our previous interpretations that while our dataset includes a small number of threat users, they are more vocal and a minority of these users tend to generate a comparatively larger amount of content than empathy users.

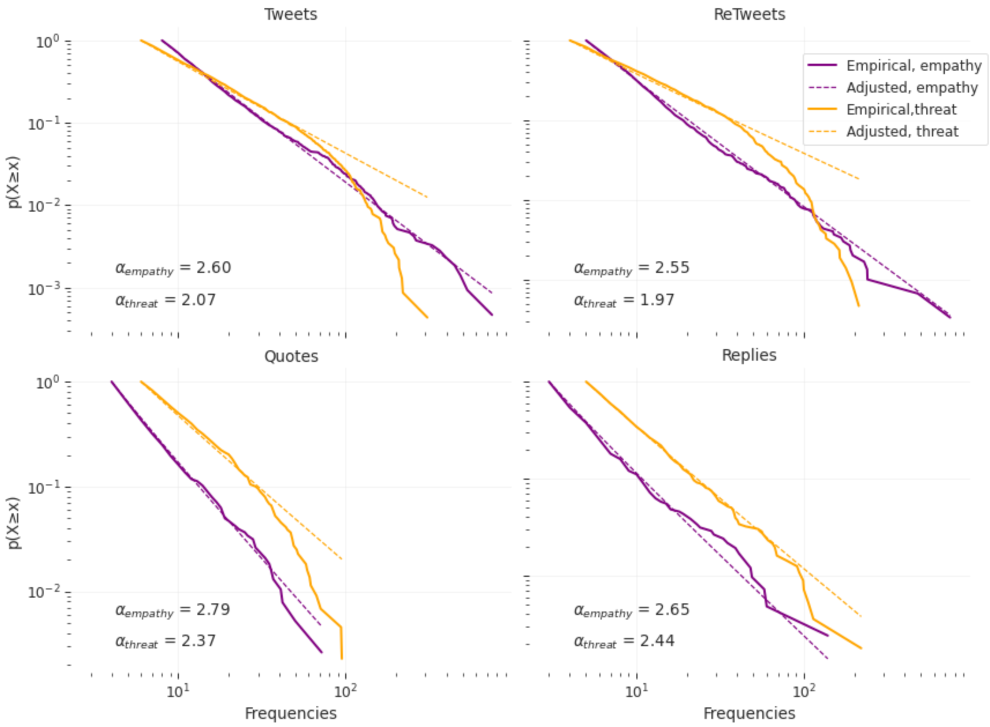


Figure 6. Powerlaw distributions for content metrics.

5.3. Psycholinguistic Characterisation

We also analysed the manifestation of cognitive and emotional structures in the text based on the LIWC lexicon and answered our question *RQ3*. Figure 7 displays the differences in the relative frequency of the LIWC categories between each group, i.e., *empathy* and *threat*.

The results reveal that categories linked to social, posemo (positive emotions), school, comm (communication), humans, optim (optimism and energy), family, affect (affective processes), incl (inclusion), and posfeel (positive feelings) are found more often in the *empathy* group. These are all psycholinguistic concepts used to describe empathetic attitudes associated with contact theory hypotheses. Conversely, categories related to motion (defined by words such as move, walk, go out), othref (reference to other people), present, negemo (negative emotions), death, inhib (inhibition), anger, money, anx (anxiety), and job are more commonly found in expressions used by the *threat* group. This is consistent with the threat theory, as immigrants can be perceived as an economic threat and labour competition.

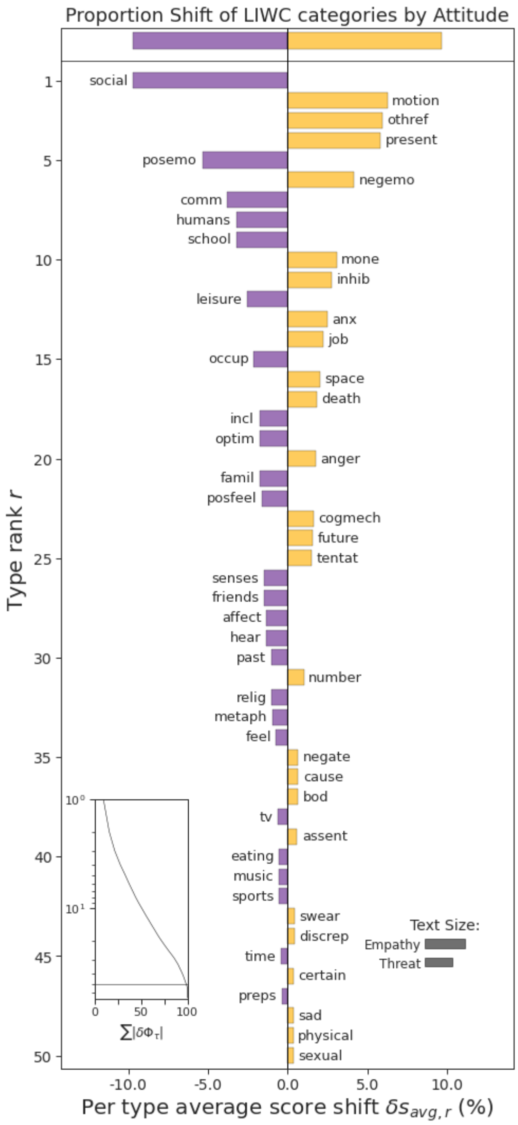


Figure 7. Proportion shift of LIWC categories in tweets grouped by *empathy* (on the left in purple) and *threat* (on the right in orange) attitude.

5.4. Dynamics and Events

Finally, we examined changes in user engagement with the migration debate over time. We estimated the daily difference in participation rate between user profile stance; that is, proportion of users in each stance group actively engaging with Twitter content about migration (see Figure 8). A positive difference indicates greater engagement among *empathy* users than *threat* users, with negative values denoting greater engagement among the latter. The resulting differences are small, though notable discrepancies exist on specific dates, particularly in the second half of 2017. Peak discrepancies in engagement coincide with key migration-related events that received mass news media coverage:

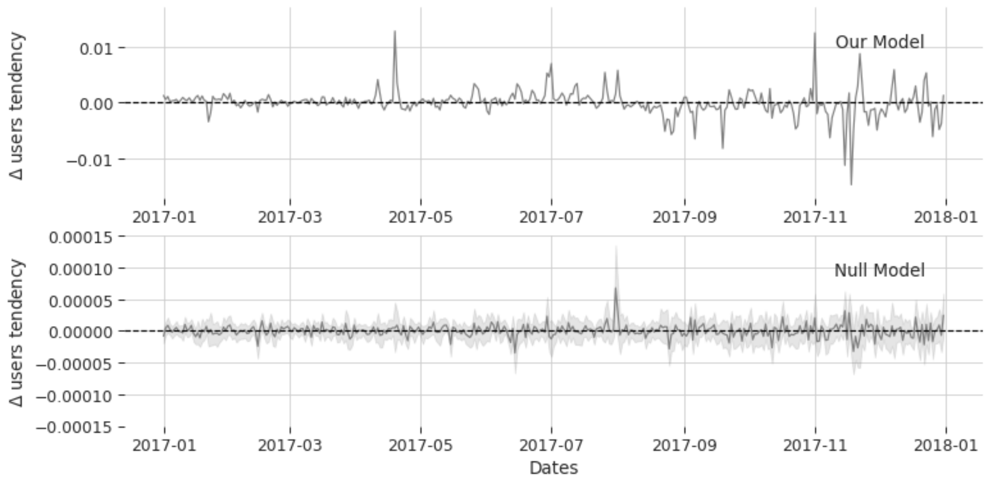


Figure 8. Daily difference in the number of users between *empathy* and *threat* groups. Top: our model, and bottom: the null model.

- **19/04:** The Chilean Census was carried out. Tweets regarding foreign interviewers went viral (positive trend spike).
- **01/07:** The National Institute of Human Rights filed a complaint against a migrant trafficking gang (positive trend spike).
- **06/09:** The Minister of Home Affairs, Mario Fernández, faced an appeal at the National Congress after a delay in the Migration Law (negative trend spike).
- **19/09:** A sign was installed in Talca (a city on central Chile) urging Haitians to join the Communist Party (negative trend spike).
- **01/11:** Director of the Central Public Hospital declared that Joane Florvil (Haitian woman who died after being arrested by Chilean police) had been beaten at the police station (positive trend spike).
- **15/11:** Senator Fulvio Rossi from Antofagasta (a city in northern Chile) was stabbed. He stated that "the attacker had a foreign accent and would be a black person" (negative trend spike).
- **18/11:** Haitian immigrants attacked police in a commercial neighbourhood in downtown Santiago (negative trend spike).
- **22/11:** (1) The Court declared the posthumous innocence of Joane Florvil. (2) Michelle Bachelet recognised not only the heroic act of Richard Joseph (Haitian citizen who rescued a woman who fell from a building) but also a set of positive human values in migration (positive trend spike).

We performed a similar analysis estimating the daily differences in the number of tweets between threat and empathy groups (see Figure 9). The results from this analysis display similar patterns of spikes as those found examining differences in user engagement. However, it highlights two key events broadcasted by the national news media: (i) Michelle Bachelet's announcement of a new visa for migrant children and youth through the "Chile welcomes you" program on 26 July (positive trend spike); and (ii) a confirmed case of an Haitian citizen with leprosy in Valdivia (a city in southern Chile) on July, 31st (negative trend spike).

We validated the differences in both measurements by comparing them with a null model, where attitudes were assigned at random (maintaining the original distribution of attitudes) in 1 K dataset permutations. Figures 8 and 9 display the null model timeseries. Given that the identified peaks are far from the 95% interval of each null model, the patterns described here are significant.

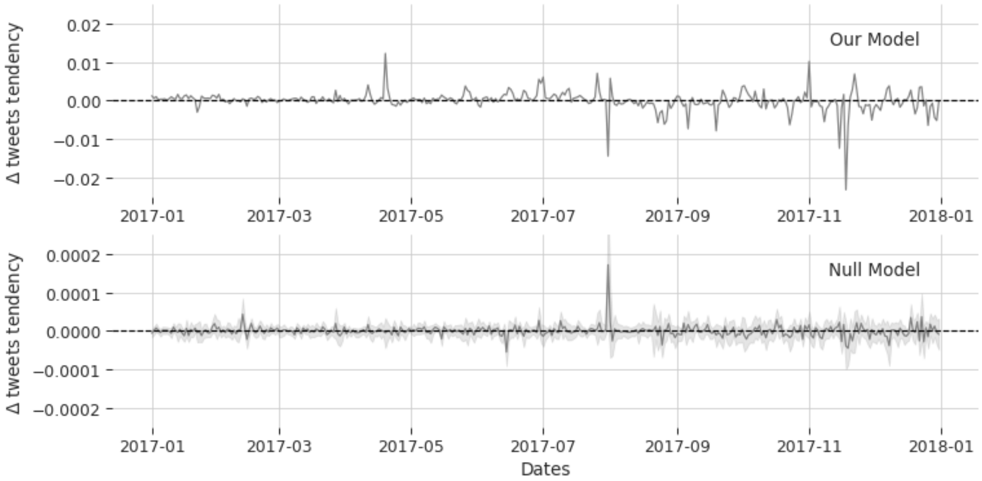


Figure 9. Daily difference in the number of tweets between *empathy* and *threat* groups. **Top**; our model, and **bottom**; the null model.

These results highlight the important role of national news media outlets in shaping the formation and expression of attitudes towards immigration in Chile. All the key events identified above were broadcasted via national television and featured on major national newspapers. As highlighted in intergroup contact and threat theories, the way in which these events are portrayed may contribute to the formation and intensified manifestation of empathetic and threatening attitudes on social media. For example, highlighted cases of physical attacks by immigrants may foment threat perceptions to public safety and hence invite expressions of negative behaviour towards immigration. The case of the Haitian citizen with leprosy is an interesting example, coinciding with intensification of tweet content generation by *threat* users and exceeding that produced by *empathy* users, despite a greater number of the latter group engaged in the discussion.

In contrast, other events portrayed by the news media were met with empathetic reactions. These include: the creation of a new visa for migrant children and young people, the case of a migrant trafficking gang, of a Haitian citizen who died while being arrested by the police, and immigrants being depicted as victims of violence and discrimination. A prominent example is also the case of a Haitian citizen who rescued a Chilean woman featuring how immigrants can contribute to the local society.

6. Discussion

6.1. Key Results

Social media is a new, dynamic and open space to express opinions and feelings, collect data and better understand public perception of immigration and offers the opportunity to overcome key limitations of traditional data sources. Social media offers the potential to monitor online public opinion about migration in near-real time at unprecedented spatial and temporal granularity, and understand the actual extent of empathy or negativity in these opinions in raw format. To unlock these potentials, we developed a novel, reproducible and open framework to measure and monitor changes in immigration sentiment. Particularly, the proposed framework enables the classification of users' stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and differences in engagement and content generation at a daily temporal scale.

We presented evidence of the composition of the Chilean immigration sentiment network on Twitter. We found that a larger share (72%) of the user network displays positive and empathetic sentiment towards immigration, and that the proportion of the network associated with negative immigration attitudes was small (25%). However, we also found that users displaying negative anti-immigration profiles tended to produce content at a significantly greater rate, producing up to 50% more content per user than users displaying positive immigration profiles. We find interesting the apparent contrast between public surveys held in Chile and our findings: reportedly, a majority of Chilean nationals agree with the following statements: "the country should take more drastic measures to exclude illegal immigrants." However, the survey is measuring something different: agreement with a statement regarding a specific group of people rather than the expression of an attitude toward a phenomenon. Furthermore, the population under study is different, as the survey samples the population aiming at national representativeness. Thus, a future line of work should be the measurement of representativeness of the Twitter population, and the definition of a methodology to compare survey results with results from our work.

We present a characterisation of the feature and psycholinguistic content of positive and negative immigration user profiles. Users with a pro-immigration profile often use neutral terms recognising migration as a process, and emotional content relating to social, positive emotions, communication, school, humans and family. By contrast, users with an anti-immigration profile tend to use terms denoting a difference between immigrants (*them*) and the native population (*us*), and emotional content relating to motion, other people, present, negative emotions, death, inhibition, anger, money, anxiety and work. We also found evidence of strong alignment between users with anti-immigration views and conservative political ideologies. Our temporal analysis also revealed pronounced peaks in daily user engagement and content generation activity in response to key migration-related events, particularly events featured in news media.

6.2. Interpretation

Our findings of a dominant base of users with pro-immigration profiles are consistent with existing prevalent trends in most of the world's regions [80]. Our findings also suggest that although the user base with anti-immigration profiles may be small, it produces and disseminates content at a significantly faster rate. Similar to the effect generated by fake news, the degree of novelty of anti-immigration content and resulting emotional reactions may be the cause of its rapid spread and generation [81].

We also showed that empathetic user profiles were linked to positive emotions, inspiring respect and unity largely in support of immigrants' human and civil rights. Anti-immigration user profiles were associated with negative emotions, calling for stricter immigration laws and claims of migrants "stealing" jobs from locals. We also found evidence of strong alignment between anti-immigration user profiles and conservative political ideologies. This pattern is not specific to Chile, but it is prevalent across industrialised countries [4].

Our analysis revealed high variability in daily user engagement and content generation activity as a result of key migration-related events. These events were prominently featured in national news media, highlighting the pivotal role that news media outlets may play in shaping the formation and expression of attitudes towards immigration in Chile. This calls for a careful approach in the way in which news media outlets portray news items involving immigrants.

6.3. Limitations and Future Work

There are two key aspects that need further exploration and that limit the scope of our results: the dynamic analysis of LIWC categories associated with attitudes, and the representativeness of Twitter. In terms of dynamics, it would be interesting to study the temporal distribution of psycholinguistic categories within empathic and threatened attitudes. This would enable quantifying the potential influence of news events on attitudes. This would provide a way to measure the effect of events and their depictions and narratives on how people feel with respect to migration.

In terms of representativeness, we acknowledge that Twitter is a biased sample of the population [82]. However, Twitter is one of the most widely used applications in Chile [83], and it reflects some of its cultural aspects, such as the centralisation of the country [84]. Furthermore, a Twitter-based analysis of the abortion discussion in Chile was found to present equal insights as those from the main national survey that covered the issue [85], hinting that there are social insights that can be derived from the platform. Thus, we propose that this work provides insights with respect to the discussion, although the representativeness of such insights is yet to be determined.

One line of work we sought to explore was to conduct a spatial analysis of attitudes, to understand the relationship between attitudes and the actual presence of immigrants in a place. This would provide a way to measure real and imagined threat attitudes [86], link virtual and physical places of expression and coexistence in a single analytical framework, and allow us to identify socio-demographic characteristics associated with the various inferred attitudes. However, as documented in the literature [10], the spatial representation of Twitter data is limited. Less than three per cent of tweets are geolocated [87]. We recognise that addressing these biases is an active area of research, and case-specific weighting schemes have been proposed to ensure the statistical and spatial representativeness of social media data [88]. However, developing weighting schemes requires knowledge of social media users' profiles. While this information can be obtained with some level of accuracy from Facebook, access to Twitter users' personal attributes is very limited. Therefore, we cannot guarantee that our results represent the general population.

7. Conclusions

In this paper, we present an analytical framework for monitoring attitudes towards immigration. Specifically the proposed framework enables the classification of users' stances of positive and negative attitudes towards immigrants and characterisation of these profiles quantitatively summarising users' content and temporal stance trends.

We applied the proposed framework to 2017 Twitter data from Chile, to capture changes in the virtual public discussion about migration during a period of a surge in immigration. We presented evidence of positive *empathetic* attitudes being expressed by a broad group of users, representing expressions of support for the immigrant community. Particularly these supportive expressions relate to calls for respect, dignity and treatment of immigrants' human and civil rights. Conversely, we provided evidence revealing that negative *threatening* attitudes towards immigration emerge from a reduced number of users, and that these attitudes are prevalent in discussions calling for stricter migrant regulation and concerns about labour competition. We also showed that negative attitudes are more commonly manifested and tend to intensify during instances of negative portrayals of immigrants. These results suggest that media news outlets play a critical role in the spread of negative representations of immigrants, and highlight the need for a more careful approach in the way in which events involving migrants are communicated. Media news outlets should consider the potential impact of misinformation fuelling misconceptions and prejudiced behaviour against immigrants. More broadly, our results demonstrate the need for a systematic approach to monitor immigration sentiment and identify shifts in attitudes towards immigrants. Such approach can enable rapid and effective mitigation plans to address misconceptions and prejudice comments against immigrants and to cushion the long-term formation of negative migration attitudes and their detrimental impacts on national social cohesion.

Author Contributions: Conceptualisation, Y.F.-V., E.G.-G. and F.R.; methodology, Y.F.-V. and E.G.-G.; software, E.G.-G.; validation, Y.F.-V., E.G.-G. and F.R.; formal analysis, Y.F.-V.; investigation, Y.F.-V., E.G.-G. and F.R.; resources, E.G.-G.; data curation, E.G.-G.; writing—original draft preparation, Y.F.-V.; writing—review and editing, Y.F.-V., E.G.-G. and F.R.; visualisation, Y.F.-V. and E.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The tweet identifiers and the classification results needed to replicate the results from this paper are available in the following shared folder: https://drive.google.com/drive/folders/1MmZyNMVtxpgcCyNhLy6UgCZjFALYi72r?usp=sharing, accessed on 16 September 2021.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

API	application programming interface
ICT	intergroup contact theory
ITT	integrated threat theory
LIWC	linguistic inquiry and word count
RQ	research question
RTs	retweets
TF-IDF	term frequency—inverse document frequency
UK	United Kingdom

References

1. European Political Strategy Centre. European Commission. 10 Trends Shaping Migration. Available online: https://op.europa.eu/s/oq7V (accessed on 16 September 2021).
2. Bail, C.A.; Argyle, L.P.; Brown, T.W.; Bumpus, J.P.; Chen, H.; Hunzaker, M.F.; Lee, J.; Mann, M.; Merhout, F.; Volfovsky, A. Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 9216–9221. [CrossRef] [PubMed]
3. Greven, T. *The Rise of Right-Wing Populism in Europe and the United States. A Comparative Perspective [La Emergencia del Populismo de Derechas en Europa y Estados Unidos. Una Perspectiva Comparada]*; Friedrich Ebert Foundation: Washington, DC, USA, 2016; pp. 1–8.
4. Dennison, J.; Geddes, A. A rising tide? The salience of immigration and the rise of anti-immigration political parties in Western Europe. *Political Q.* **2019**, *90*, 107–116. [CrossRef]
5. Blinder, S.; Allen, W. UK public opinion toward immigration: Overall attitudes and level of concern. In *Migration Observatory Briefing, COMPAS*; University of Oxford: Oxford, UK, 2016.
6. Cheong, P.H.; Edwards, R.; Goulbourne, H.; Solomos, J. Immigration, social cohesion and social capital: A critical review. *Crit. Soc. Policy* **2007**, *27*, 24–49. [CrossRef]
7. Penninx, R.; Spencer, D.; Van Hear, N. *Migration and Integration in Europe: The State of Research*; Economic and Social Research Council: Swindon, UK, 2008.
8. United Nations. Sustainable Development Goals. Available online: https://www.un.org/sustainabledevelopment/sustainable-development-goals/ (accessed on 16 September 2021).
9. Global Compact for Migration. Global Compact for Safe, Orderly and Regular Migration. Available online: https://refugeesmigrants.un.org/sites/default/files/180713_agreed_outcome_global_compact_for_migration.pdf (accessed on 16 September 2021).
10. Rowe, F.; Mahony, M.; Graells-Garrido, E.; Rango, M.; Sievers, N. *Using Twitter Data to Monitor Immigration Sentiment. Practitioners' Guidebook*; International Organization for Migration; United Nations: Berlin, Germany, 2021.
11. Grigorieff, A.; Roth, C.; Ubfal, D. Does information change attitudes toward immigrants? *Demography* **2020**, *57*, 1117–1143. [CrossRef]
12. Dennison, J.; Dražanová, L. *Public Attitudes on Migration: Rethinking How People Perceive Migration: An Analysis of Existing Opinion Polls in the Euro-Mediterranean Region*; Technical Report; European University Institute: Fiesole, Italy, 2018.
13. Burns, P.; Gimpel, J.G. Economic insecurity, prejudicial stereotypes, and public opinion on immigration policy. *Political Sci. Q.* **2000**, *115*, 201–225. [CrossRef]
14. Rowe, F.; Mahony, M.; Graells-Garrido, E.; Rango, M.; Sievers, N. Using Twitter to track immigration sentiment during early stages of the COVID-19 pandemic. *SocArXiv* **2021**. [CrossRef]
15. Wu, S.; Hofman, J.M.; Mason, W.A.; Watts, D.J. Who says what to whom on Twitter. In Proceedings of the 20th International Conference on World Wide Web, Yderabad, India, 28 March–1 April 2011; pp. 705–714.
16. Garimella, K.; De Francisci Morales, G.; Gionis, A.; Mathioudakis, M. The effect of collective attention on controversial debates on social media. In Proceedings of the 2017 ACM on Web Science Conference, Troy, NY, USA, 25–28 June 2017; pp. 43–52.
17. Graells-Garrido, E.; Baeza-Yates, R.; Lalmas, M. Every colour you are: Stance prediction and turnaround in controversial issues. In Proceedings of the 12th ACM Conference on Web Science, Southampton, UK, 6–10 July 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 174–183. [CrossRef]

18. Bosco, C.; Patti, V.; Bogetti, M.; Conoscenti, M.; Ruffo, G.F.; Schifanella, R.; Stranisci, M. Tools and Resources for Detecting Hate and Prejudice against Immigrants in Social Media. Symposium III. Social Interactions in Complex Intelligent Systems (SICIS). Artificial Intelligence and Simulation of Behaviour (AISB). 2017; pp. 79–84. Available online: https://iris.unito.it/retrieve/handle/2318/1637776/332903/paperBATH.pdf (accessed on 16 September 2021).
19. Sanguinetti, M.; Poletto, F.; Bosco, C.; Patti, V.; Stranisci, M. An Italian Twitter Corpus of Hate Speech against Immigrants. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, 7–12 May 2018. Available online: https://www.aclweb.org/anthology/L18-1443 (accessed on 16 September 2021).
20. Basile, V.; Bosco, C.; Fersini, E.; Debora, N.; Patti, V.; Pardo, F.M.R.; Rosso, P.; Sanguinetti, M. Semeval-2019 task 5: Multilingual detection of hate speech against immigrants and women in Twitter. In Proceedings of the 13th International Workshop on Semantic Evaluation. Association for Computational Linguistics, Minneapolis, MN, USA, 6–7 June 2019; pp. 54–63. Available online: https://arxiv.org/abs/2011.13238 (accessed on 16 September 2021).
21. Comandini, G.; Patti, V. An Impossible Dialogue! Nominal Utterances and Populist Rhetoric in an Italian Twitter Corpus of Hate Speech against Immigrants. In Proceedings of the Third Workshop on Abusive Language Online, Florence, Italy, 1 August 2019; pp. 163–171. Available online: https://www.aclweb.org/anthology/W19-3518/ (accessed on 16 September 2021)
22. Calderón, C.A.; de la Vega, G.; Herrero, D.B. Topic modeling and characterization of hate speech against immigrants on Twitter around the emergence of a far-right party in Spain. *Soc. Sci.* **2020**, *9*, 188. [CrossRef]
23. Porcher, S.; Renault, T. Social distancing beliefs and human mobility: Evidence from Twitter. *PLoS ONE* **2021**, *16*, e0246949. [CrossRef]
24. Waseem, Z.; Hovy, D. Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. In Proceedings of the NAACL Student Research Workshop, San Diego, CA, USA, 12–17 June 2016; pp. 88–93. [CrossRef]
25. Bovet, A.; Makse, H.A. Influence of fake news in Twitter during the 2016 US presidential election. *Nat. Commun.* **2019**, *10*, 1–14. [CrossRef]
26. Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; Lazer, D. Fake news on Twitter during the 2016 US presidential election. *Science* **2019**, *363*, 374–378. [CrossRef]
27. Rowe, F.; Darlington-Pollock, F. *Quantifying and Understanding the Extent of Residential Segregation of Recent Immigrants in Chile*; Conference Paper; The British Society for Population Studies: Cardiff, UK, 2019.
28. Allport, G.W.; Clark, K.; Pettigrew, T. *The Nature of Prejudice*; Addison-Wesley Reading: Boston, MA, USA, 1954.
29. Nelson, T.D. *Handbook of Prejudice, Stereotyping, and Discrimination*; Psychology Press: Hove, UK, 2009.
30. Stephan, W.G.; Stephan, C.W. An integrated threat theory of prejudice. *Reducing Prejudice and Discrimination*; Psychology Press: Hove, UK, 2000; pp. 23–45.
31. Paluck, E.L.; Green, S.A.; Green, D.P. The contact hypothesis re-evaluated. *Behav. Public Policy* **2019**, *3*, 129–158. [CrossRef]
32. Pettigrew, T.F.; Tropp, L.R. How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *Eur. J. Soc. Psychol.* **2008**, *38*, 922–934. [CrossRef]
33. Pettigrew, T.F.; Tropp, L.R. A meta-analytic test of intergroup contact theory. *J. Personal. Soc. Psychol.* **2006**, *90*, 751. [CrossRef] [PubMed]
34. Barlow, F.K.; Louis, W.R.; Hewstone, M. Rejected! Cognitions of rejection and intergroup anxiety as mediators of the impact of cross-group friendships on prejudice. *Br. J. Soc. Psychol.* **2009**, *48*, 389–405. [CrossRef]
35. Stephan, W.G.; Stephan, C.W. Intergroup anxiety. *J. Soc. Issues* **1985**, *41*, 157–175. [CrossRef]
36. Bertrand, M.; Duflo, E. Field experiments on discrimination. *Handb. Econ. Field Exp.* **2017**, *1*, 309–393.
37. Stephan, W.G.; Finlay, K. The role of empathy in improving intergroup relations. *J. Soc. Issues* **1999**, *55*, 729–743. [CrossRef]
38. Esses, V.M.; Dovidio, J.F.; Jackson, L.M.; Armstrong, T.L. The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *J. Soc. Issues* **2001**, *57*, 389–412. [CrossRef]
39. Ha, S.E. The consequences of multiracial contexts on public attitudes toward immigration. *Political Res. Q.* **2010**, *63*, 29–42. [CrossRef]
40. Hanson, G.H.; Scheve, K.; Slaughter, M.J. Public finance and individual preferences over globalization strategies. *Econ. Politics* **2007**, *19*, 1–33. [CrossRef]
41. International Organization for Migration (OIM). DATA BULLETIN SERIES. Informing the Implementation of the Global Compact for Migration. Available online: https://publications.iom.int/system/files/pdf/gmdacbulletins.pdf (accessed on 16 September 2021).
42. Sylwester, K.; Purver, M. Twitter language use reflects psychological differences between democrats and republicans. *PLoS ONE* **2015**, *10*, e0137422.
43. Golder, S.A.; Macy, M.W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* **2011**, *333*, 1878–1881. [CrossRef]
44. Ritter, R.S.; Preston, J.L.; Hernandez, I. Happy tweets: Christians are happier, more socially connected, and less analytical than atheists on Twitter. *Soc. Psychol. Personal. Sci.* **2014**, *5*, 243–249. [CrossRef]
45. Conover, M.D.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Menczer, F.; Flammini, A. Political polarization on Twitter. In Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, 17–21 July 2011.
46. Garcia-Gavilanes, R.; Quercia, D.; Jaimes, A. Cultural dimensions in Twitter: Time, Individualism and Power. In Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, Cambridge, MA, USA, 8–11 July 2013.
47. Darwish, K.; Magdy, W.; Rahimi, A.; Baldwin, T.; Abokhodair, N. Predicting online islamophopic behavior after# parisattacks. *J. Web Sci.* **2018**, *4*, 34–52.

48. Lamanna, F.; Lenormand, M.; Salas-Olmedo, M.H.; Romanillos, G.; Gonçalves, B.; Ramasco, J.J. Immigrant community integration in world cities. *PLoS ONE* **2018**, *13*, e0191612. [CrossRef] [PubMed]
49. Fortuna, P.; Nunes, S. A survey on automatic detection of hate speech in text. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–30. [CrossRef]
50. Ribeiro, M.; Calais, P.; Santos, Y.; Almeida, V.; Meira Jr, W. Characterizing and Detecting Hateful Users on Twitter. In Proceedings of the International AAAI Conference on Web and Social Media, Stanford, CA, USA, 25–28 June 2018; Volume 12. Available online: https://arxiv.org/abs/1803.08977v1 (accessed on 16 September 2021).
51. Zhang, Q.; Liang, S.; Lipani, A.; Ren, Z.; Yilmaz, E. From Stances' Imbalance to Their Hierarchical Representation and Detection. In Proceedings of the The World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 2323–2332. [CrossRef]
52. Garimella, K.; De Francisci Morales, G.; Gionis, A.; Mathioudakis, M. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In Proceedings of the 2018 World Wide Web Conference on World Wide Web, Lyon, France, 23–27 April 2018; International World Wide Web Conferences Steering Committee: Lyon, France, 2018; pp. 913–922.
53. Garimella, K.; Morales, G.D.F.; Gionis, A.; Mathioudakis, M. Quantifying controversy on social media. *ACM Trans. Soc. Comput.* **2018**, *1*, 3. [CrossRef]
54. Conover, M.D.; Gonçalves, B.; Ratkiewicz, J.; Flammini, A.; Menczer, F. Predicting the political alignment of Twitter users. In Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9–11 October 2011; IEEE: Boston, MA, USA, 2011; pp. 192–199.
55. Lu, H.; Caverlee, J.; Niu, W. Biaswatch: A lightweight system for discovering and tracking topic-sensitive opinion bias in social media. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, 19–23 October 2015; ACM: Melbourne, Australia, 2015; pp. 213–222.
56. Wilson, T.; Wiebe, J.; Hoffmann, P. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Vancouver, BC, Canada, 6–8 October 2005; pp. 347–354.
57. Pang, B.; Lee, L. *Opinion Mining and Sentiment Analysis*; Association for Computing Machinery: New York, NY, USA, 2008, Volume 2, pp. 1–135. [CrossRef]
58. Pennebaker, J.W.; Mehl, M.R.; Niederhoffer, K.G. Psychological aspects of natural language use: Our words, our selves. *Annu. Rev. Psychol.* **2003**, *54*, 547–577. [CrossRef]
59. De Choudhury, M.; Counts, S.; Horvitz, E.J.; Hoff, A. Characterizing and predicting postpartum depression from shared Facebook data. In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, Baltimore, MA, USA, 15–19 February 2014; pp. 626–638.
60. Harman, G.; Dredze, M.H. Measuring post traumatic stress disorder in Twitter. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, ICWSM, Palo Alto, CA, USA, 1–4 June 2014. Available online: https://www.cs.jhu.edu/~mdredze/publications/2014_icwsm_ptsd.pdf (accessed on 16 September 2021).
61. Tumasjan, A.; Sprenger, T.O.; Sandner, P.G.; Welpe, I.M. Predicting elections with Twitter: What 140 characters reveal about political sentiment. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010.
62. Vasquez-Henriquez, P.; Graells-Garrido, E.; Caro, D. Tweets on the go: Gender differences in transport perception and its discussion on social media. *Sustainability* **2020**, *12*, 5405. [CrossRef]
63. Ramírez-Esparza, N.; Pennebaker, J.W.; García, F.A.; Suriá, R. La psicología del uso de las palabras: Un programa de computadora que analiza textos en español. *Rev. Mex. Psicol.* **2007**, *24*, 85–99.
64. Economic Commission for Latin America and the Caribbean (ECLAC). Demographic Observatory of Latin America 2018: International Migration. Available online: https://www.oecd.org/acerca/miembros-y-socios/ (accessed on 16 September 2021).
65. Dammert, L.; Erlandsen, M. Migración, miedos y medios en la elección presidencial en Chile (2017). *Revista CS* **2020**, *31*, 43–76. [CrossRef]
66. Batarce, C. La Tercera. Ciudadano haitiano en Valdivia es Sospechoso de Padecer Lepra. Available online: http://www2.latercera.com/noticia/haitiano-se-convierte-primer-caso-lepra-chile-continental/# (accessed on 16 September 2021).
67. Public Studies Center. Estudio Nacional de Opinión Pública N°79, Abril-Mayo 2017. Available online: https://www.cepchile.cl/cep/encuestas-cep/encuestas-2010-2019/estudio-nacional-de-opinion-publica-abril-mayo-2017 (accessed on 16 September 2021).
68. Graells-Garrido, E.; Lalmas, M.; Baeza-Yates, R. Encouraging Diversity-and Representation-Awareness in Geographically Centralized Content. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016; pp. 7–18.
69. Bryden, J.; Funk, S.; Jansen, V.A. Word usage mirrors community structure in the online social network Twitter. *EPJ Data Sci.* **2013**, *2*, 1–9. [CrossRef]
70. Nerghes, A.; Lee, J.S. The refugee/migrant crisis dichotomy on Twitter: A network and sentiment perspective. In Proceedings of the 10th ACM Conference on Web Science, Amsterdam, The Netherlands, 27–30 May 2018; pp. 271–280.
71. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

72. Manivannan, A.; Yow, W.Q.; Bouffanais, R.; Barrat, A. Are the different layers of a social network conveying the same information? *EPJ Data Sci.* **2018**, *7*, 34. [CrossRef]
73. Monroe, B.L.; Colaresi, M.P.; Quinn, K.M. Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Anal.* **2008**, *16*, 372–403. [CrossRef]
74. Serrano, M.Á.; Flammini, A.; Menczer, F. Modeling statistical properties of written text. *PLoS ONE* **2009**, *4*, e5372. [CrossRef] [PubMed]
75. Kwak, H.; Lee, C.; Park, H.; Moon, S. What is Twitter, a social network or a news media? In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 591–600.
76. Pennebaker, J.W.; Francis, M.E.; Booth, R.J. Linguistic inquiry and word count: LIWC 2001. *Mahway Lawrence Erlbaum Assoc.* **2001**, *71*, 2001.
77. Gallagher, R.J.; Frank, M.R.; Mitchell, L.; Schwartz, A.J.; Reagan, A.J.; Danforth, C.M.; Dodds, P.S. Generalized word shift graphs: A method for visualizing and explaining pairwise comparisons between texts. *arXiv* **2020**, arXiv:2008.02250.
78. Graells-Garrido, E.; Lalmas, M.; Baeza-Yates, R. Data portraits and intermediary topics: Encouraging exploration of politically diverse profiles. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016; pp. 228–240.
79. Lu, Y.; Zhang, P.; Cao, Y.; Hu, Y.; Guo, L. On the frequency distribution of retweets. *Procedia Comput. Sci.* **2014**, *31*, 747–753. [CrossRef]
80. International Organization for Migration. How the World Views Migration, Geneva. 2015 Available online: https://publications.iom.int/system/files/how_the_world_gallup.pdf (accessed on 16 September 2021).
81. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* **2018**, *359*, 1146–1151. [CrossRef]
82. Baeza-Yates, R. Bias on the web. *Commun. ACM* **2018**, *61*, 54–61. [CrossRef]
83. Graells-Garrido, E.; Caro, D.; Miranda, O.; Schifanella, R.; Peredo, O.F. The WWW (and an H) of mobile application usage in the city: The what, where, when, and how. In Proceedings of the Companion of the The Web Conference 2018 on The Web Conference 2018, International World Wide Web Conferences Steering Committee, Lyon, France, 23–27 April 2018; pp. 1221–1229.
84. Graells-Garrido, E.; Lalmas, M. Balancing diversity to counter-measure geographical centralization in microblogging platforms. In Proceedings of the 25th ACM Conference on Hypertext and Social Media, Santiago, Chile, 1–4 September 2014; pp. 231–236.
85. Graells-Garrido, E.; Baeza-Yates, R.; Lalmas, M. How representative is an abortion debate on Twitter? In Proceedings of the 10th ACM Conference on Web Science, Boston, MA, USA, 30 June–3 July 2019; pp. 133–134.
86. Kopstein, J.S.; Wittenberg, J. Does familiarity breed contempt? Inter-ethnic contact and support for illiberal parties. *J. Politics* **2009**, *71*, 414–428. [CrossRef]
87. Twitter. Tutorials: Tweet GeospatialMetadata. 2021. Available online: https://developer.twitter.com/en/docs/tutorials/tweet-geo-metadata (accessed on 16 September 2021).
88. Grow, A.; Perrotta, D.; Del Fava, E.; Cimentada, J.; Rampazzo, F.; Gil-Clavel, S.; Zagheni, E.; Flores, R.D.; Ventura, I.; Weber, I.; et al. *How Reliable Is Facebook's Advertising Data for Use in Social Science Research? Insights from a Cross-National Online Survey*; Technical Report; Max Planck Institute for Demographic Research: Rostock, Germany, 2021.