Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed
Computing

Analysis of Protests in Nigeria using Social Media Data

Master's Thesis

Katharina Börsig
kboersig@ethz.ch

Distributed Computing Group
Computer Engineering and Networks Laboratory
ETH Zürich

International Conflict Research Group
Center for Comparative and International Studies (CIS)
ETH Zürich

Supervisors:
Zhao Meng, Dr. Yannick Pengl
Prof. Dr. Roger Wattenhofer, Prof. Dr. Lars-Erik Cederman

July 4, 2022

Acknowledgements

First, I would like to thank my first supervisor Zhao Meng for supporting my work from the computer science perspective and discussing various technical approaches and ideas. Further, I would like to thank my second supervisor Yannick Pengl for supporting my project from the political science perspective and discussing implications of analysis results in the context of Nigeria.

Lastly, I would like to thank my family and friends for supporting me throughout this project.

Abstract

Comparative studies of the relation between online social movements and offline collective action focus on case studies leaving a large amount of available social media data unused. We take the availability of social media data as an opportunity to gain a deeper understanding of the relation between online social movements and offline collective actions in Nigeria from a quantitative perspective using Big Data analysis techniques. In order to determine how political discourse shapes the Nigerian Twitter landscape, we explore topics modeled on Nigerian Twitter data. The exploratory analysis indicates the importance of political discourse on Nigerian Twitter and validates its suitability for protest analysis. Further, we introduce an AI-based approach for automatically identifying and mapping Nigerian online social movements to protests in Nigeria. Quality assessment of our mapping shows that it successfully integrates social media and protest data. Lastly, we perform quantitative comparative analysis across hundreds of relations between online social movements and protests using NLP and network analysis, among others. We identify relevant characteristics and their typical behaviors that determine the relationship between online social movements and protests in Nigeria.

Contents

Chapter 1
Introduction

Over the last decade, social media has played an essential role in various social debates, including the Arab Spring [1], the London riots [2], and the Occupy Wall Street [3] movements. Hence, the power of social media in the context of politics is not to be underestimated. While political discourse on social media is already on the agenda in most developed countries, internet penetration and, thereby also, social media is still on the rise in many African countries [4]. Only recently, in 2021, Twitter announced to establish its first headquarter on the African continent [5].

With the rise of social media come debates about social media regulations, bills, and bans. In Nigeria, President Muhammadu Buhari announced a permanent ban of Twitter on June 5th, 2021, after a questionable tweet of his was deleted two days earlier [6]. Twitter has long been a thorn in the side of the Nigerian government [7]. The ongoing '#EndSars' protests against police brutality began on Twitter [8] and escalated in 2020, leading to mass protests all over the country [9]. While the government tried to regulate social media before [10], all attempts to pass an anti-social media bill failed due to a massive outcry on social media [10]. Moreover, the Nigerian minister of information called Twitter suspicious, referring to its influence on the '#EndSars' protests [11] only a few days before the ban. In January 2022, after seven months of negotiation, the Nigerian government lifted the ban after Twitter made various concessions such as establishing a legal entity in Nigeria [12] and therefore increasing control of the Nigerian government. These actions show how intimidated the Nigerian government is by the massive usage of Twitter in the context of successful social movements and mass protests.

The relationship between online social movements and offline collective actions has been the subject of research since the rise of social media. Researchers typically perform case studies, i.e., they analyse aspects of the behavior of one large online social movement in the context of the related mass protests such as the '#BlackLivesMatter' movement [13]. The analysed aspects of the behavior and the respective analysis vary ranging from meta-data analysis [13, 14, 15] over social network analysis [3, 16, 17, 18], to Natural Language Processing [19, 20, 17]. Comparative case studies qualitatively investigated differences between social movements, mainly focusing on changes from conventional social movements to online social movements [21, 22] in the same country. Due to their nature, these case studies serve as evidence for hypotheses but don not shed light on the general transferability of research findings to other protests.

In this thesis, we use Nigeria's Twitter ban and the implicit rejection of Twitter due to its influence on social movements as an occasion to investigate the relation between online social movements and offline collective actions in Nigeria in detail. We perform a quantitative comparative analysis across hundreds of online social movements and related protests to gain insights into their relationships transferable to other protests in Nigeria. We approach this task from a computer science perspective by employing various Big Data analysis techniques, including data mining from Twitter, data integration with an AI-based mapping of tweets to protests, and data analysis using Natural Language Processing, social network analysis, and various statistical methods.

1.1 Background Information on Nigeria

Nigeria is a former British colony that became formally independent in 1960. The independence was followed by civil war and military dictatorships until democracy was established in 1999. Similar to other post-colonial African countries, Nigeria is a multinational state consisting of more than 250 ethnic groups with various cultures, religions, and languages. For facilitating linguistic unity, English was chosen as the official language [23]. Religious affiliation among the Nigerian population is roughly divided in half between Christianity and Islam. This division goes hand in hand with the prevailing north-south divide: The north of Nigeria is mainly Muslim populated, significantly poorer, less developed, and more frequently afflicted by violent attacks, including banditry by local criminal gangs and attacks by terrorist organizations such as Boko Haram.

Regarding the political system, Nigeria introduced a federal republic system based on the model of the United States with 36 states and a federal capital territory. The president holds executive power and is both head of state and head of government. Currently, Muhammadu Buhari holds the office of the Nigerian president. He was first elected in 2015 and re-elected in 2019.

Nigeria is the most populous African country and generates the largest African nominal GDP [24]. Nonetheless, it ranks low in the human development index (HDI) [25], is one of the most corrupt countries (CPI) [26], and ranks lowest in the World Internal Security & Police Index (WISPI) [27]. Further, its' human rights records remain poor (HRMI) [28]. These factors give rise to security crises in nearly every part of the country. In the northeast of Nigeria, the Islamist group Boko Haram has conducted attacks leading to over 30k fatalities so far [29]. Further, in the northwest of Nigeria, a large amount of banditry takes place, including kidnapping for ransom and cattle rustling [30]. In addition, in the southeast of Nigeria, the Indigenous People of Biafra (IPOB) fight against the military for independence [31]. Lastly, the north center of Nigeria, called the

Middle Belt states, has been the scene of the long-lasting herder-farmer conflicts between Fulani pastoralists and non-Fulani farmers[32].

As an answer to increasing banditry, the Nigerian government formed a Special Anti-Robbery Squad (SARS) in 1992. The main task of this squad is to tackle violent gang crimes such as armed robbery and kidnapping. However, the special force has been an object of criticism for multiple years by the public and various organizations, including Amnesty International [33]. They accuse SARS of a variety of human rights violations such as harassment, extortion, rape, torture, and extrajudicial killings [33]. Further, they summarize their actions as systemic human rights violations perpetrated with impunity [34].

This level of violence, security threats, and impunity has led to the Nigerian population speaking up against violence and irregularities in the government. Over the last decade, Nigerians successfully organized mass protests that gained international attention, increasing the pressure on the Nigerian government to take action. For instance, as a reaction to the kidnapping of over 250 schoolgirls in Chibok by Boko Haram, a Twitter campaign with the hashtag #BringBackOurGirls led to large-scale protests all over the world [35]. The escalating killings by bandits in Zamfara state were countered by youth protests organized by the 'North is bleeding' social media campaign [36]. Further, to oppose the human rights violations by SARS, a Twitter campaign using the hashtag #EndSARS evolved into a large-scale social movement that led to a series of mass protests in Nigeria [9]. In addition, government irregularities, including corruption or attempts to restrict freedom of expression, were countered by social media campaigns and mass protests such as the #RevolutionNow and the #AntiSocialMediaBill movement. Latter ended successfully in the suspension of the bill due to massive criticism from Nigerian and international media [10].

1.2 Research Questions

Although online social movements have been analyzed in case studies for various mass protests, quantitative analysis of online social movements on the African continent tends to be underrepresented. Further, comparative analysis has been limited to comparative case studies of typically only two prominent online social movements.

This thesis' main objective is to transition from studying individual social movements to performing quantitative analyses using Big Data techniques. In the context of Nigeria, we aim to shed light on the quantitative nature of relations between online social movements and offline collective actions to develop a deeper understanding of this relation.

Approaching this objective requires multiple steps of preparation. First, the Nigerian Twitter landscape needs to be suitable for this quantitative analysis. Hence, the first research question refers to an NLP-based exploratory analysis of the Nigerian Twitter landscape:

I) To which extent do politics and collective actions shape the Nigerian Twitter landscape?

Further, quantitative analysis requires a large number of social movements mapped to their related protests. Therefore, the second research question targets the requirement of data integration, i.e., automatic mapping of online social movements and offline collective actions:

II) How can online social movements be identified and automatically mapped to related protests?

Based on the preceding steps, the main objective can be approached. The last research question refers to the quantitative comparative analysis of the relations between online social movements and offline collective actions using NLP and network analysis methods, among others:

III) What is the nature of relations between online social movements and offline collective actions?

1.3 Outline

Besides this introduction, the thesis is organized into five chapters. In Chapter 2 we explain the technical background and elaborate on prior related work. Further, we introduce the methodology in Chapter 3 including a description of the employed data sets as well as a delineation of the experiments. In Chapter 4, we present the results of our experiments. Subsequently, we discuss the results, reflect on the limitations of our work and propose directions for future research in Chapter 5. Finally, we conclude the thesis by summarizing our findings and the answers to the proposed research questions in Chapter 6.

Chapter 2

Related Work

This chapter provides a detailed introduction to the relevant topics for this thesis and discusses prior related work. First, Section 2.1 elaborates on the field of Natural Language Processing. Second, in Section 2.2 advances in topic modeling are discussed. Finally, Section 2.3 introduces the research field of Social Media Analysis.

2.1 Natural Language Processing (NLP)

Natural Language Processing is a computerized approach that combines multiple research fields such as linguistics, psychology, and computer science. It aims to understand textual content to solve given tasks, including translation, classification, and question answering. Traditionally this field was dominated by simple statistical models such as Support Vector Machines (SVM) and Logistic Regression. With advancements in neural network architecture, AI models started to show superior performance in various NLP tasks, leading up to recent successes of deep learning models such as BERT and its variations.

Most NLP approaches follow a given schema: Raw text is transformed into numerical feature vectors, called text representations or embeddings, on which learning models are then trained for a specific task. This chapter elaborates on the evolution and the state-of-the-art of network architectures and text representations to contextualize methods applied in this thesis. In Section 2.1.1 the evolution of network architectures is discussed. Subsequently, Section 2.1.2 elaborates the evolution of text representations. Last, Section 2.1.3 explains the deep learning model BERT in more detail.

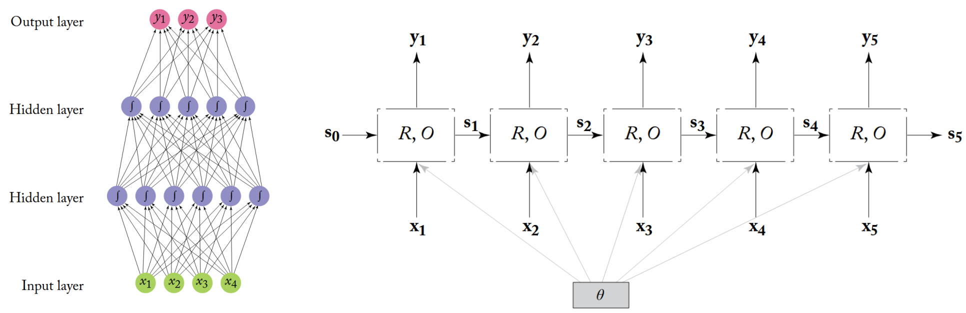
2.1.1 Neural network architectures

The architecture of neural networks has evolved significantly over the last decades from simple feed-forward neural networks to complex transformer networks. In the following, this evolution of network architecture is briefly discussed.

Feed-forward neural networks. Feed-forward neural networks (FFN) are also called multi-layer perceptrons [37], since multiple sequential layers of perceptrons

[38] aim to find non-linear patterns in data. Figure 2.1a shows a typical FFN architecture which consists of one input layer, multiple hidden layers, and one output layer. To each node in the hidden and output layer, a weighted activation function is applied, such as sigmoid or tanh function [39], which transforms the input into non-linear output. The weights are modified during training based on the model's loss, which captures the discordance between the model prediction and the true value.

Recurrent neural networks. The simple feed-forward architecture is not able to consider the order of input data which is a disadvantage when dealing with sequential data such as sentences. Since different order of words in a sentence can change its meaning, taking the word order into account is crucial for NLP tasks. Hence, recurrent neural networks (RNN)[40] were developed, which conserve the activation values of prior data points by consecutively connecting nodes of hidden layers as shown in Figure 2.1b. A variation of the RNN architecture is the long short-term memory architecture (LSTM) [41]. Every hidden node includes input, forget, and output gates that determine which information from preceding nodes should be considered to compute the next hidden state.



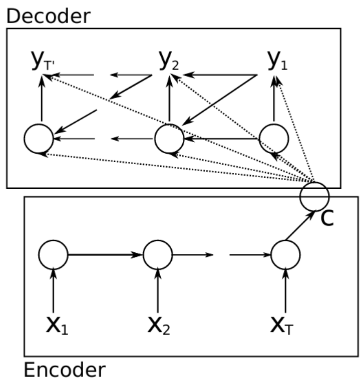
(a) Fully connected feed-forward network

(b) Recurrent Neural Network

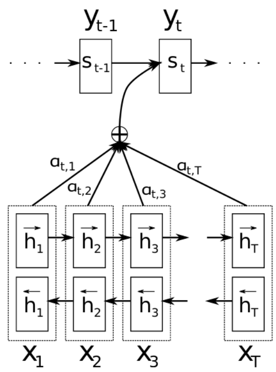
Figure 2.1: Two basic neural network architectures [39].

Encoder-Decoder. In encoder-decoder architectures [42] two RNNs are combined into one model as shown in Figure 2.2a. One RNN encodes the input sequence into a vector representation of fixed length while the other decodes this representation into an output, such as another sequence for sequence-to-sequence NLP tasks (e.g., machine translation). Since the encoder and decoder are jointly trained, the vector representation is also optimized for the given task. A drawback of this basic encoder-decoder architecture is that it cannot remember longer sentences. The compression of a large amount of information into a fixed-length vector representation is difficult when the model does not know what to pay attention to [43].

In order to bypass this drawback, an attention mechanism [43] was introduced



(a) Encoder-Decoder architecture [42].



(b) Bahdenau attention mechanism [43].

Figure 2.2: Encoder-Decoder Components: Schematic architecture and attention mechanism.

that enables the model to pay attention to specific parts of the input sequence. By using an attention mechanism, the model can search itself for positions in the sequence that are most relevant for the given task. An additive attention layer, as shown in Figure 2.2b allows the decoder to use the output of each hidden state rather than summarizing the input sequence into one single vector. Qualitative analysis of attention weights produced by a textual entailment model showed that this attention mechanism successfully pays attention to semantically coherent or contradicting words [44], similar to human intuition.

Transformers. The status-quo of language models was challenged by the introduction of transformer networks [45]. This architecture, shown in Figure 2.3, follows an encoder-decoder architecture relying only on the attention mechanism in combination with traditional feed-forward neural networks rather than applying RNN encoder and decoders.

Typically, an additive attention layer referred to as encoder-decoder attention is applied to the decoder. Further attention layers are applied for both the encoder and decoder, referred to as multi-head self-attention layers. Self-attention is a form of attention located within an encoder or decoder rather than in between. Thus, for any individual unit of the sequence, such as a word, this mechanism identifies relevant units of the same input, such as its context words, as shown in Figure 2.4a. Multi-headed self attention allows multiple self-attentions to run in parallel (see Figure 2.4b). Each attention head can capture a different property which ensures that also long-distance dependencies are considered [45]. This transformer architecture was applied successfully in various language models such as GPT-2 [46] and BERT [47].

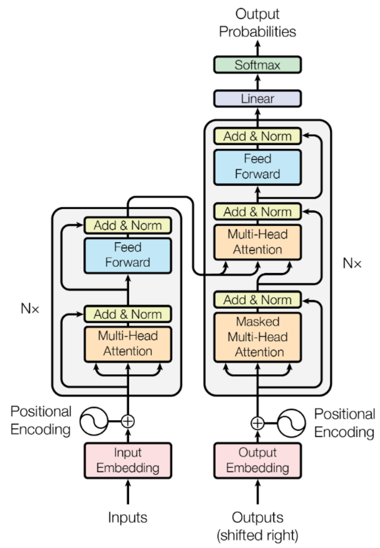
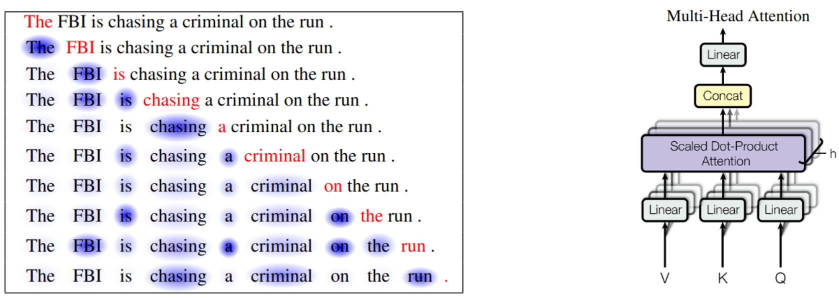


Figure 2.3: Transformer Architecture [45].



(a) Self-attention [48]

(b) Multi-head self-attention [45]

Figure 2.4: Attention mechanism applied for Transformers.

2.1.2 Text representations

In order to apply learning models to text data for NLP tasks, the text is represented numerically. Similar to network architectures, text representation approaches evolved from simple count-based to complex prediction-based techniques. In this Section, notable representations from both approaches are elaborated.

Count-based representations. The simplest count-based representation, referred to as the Bag of Words (BoW) method, determines word vectors based on the frequency of word occurrences in documents. Originally, this approach was applied for computing document similarity, and thus, the word counts represented document features [49]. However, by using the term frequencies over all documents, also word representations can be computed.

More advanced methods such as TF-IDF [50] enhance these simple representations by weighing the terms. TF-IDF multiplies the term frequency (TF) in a document with the inverse document frequency (IDF) determined by the number of documents containing this term. Thus, unique words are assigned higher values in comparison to a simple BoW representation. Nonetheless, both methods are unable to capture the context of words. By adding n-grams to the vocabulary, the context can be incorporated to some extent.

In order to enhance context capturing, co-occurrence matrices of the words can be considered. However, since co-occurrence matrices are sparse, dimension reduction techniques are required for further efficient computations. This approach is successfully applied in GloVe [51] which combines the sparse co-occurrence matrices efficiently with least-squares optimization.

Prediction-based representation. Prediction-based representations employ a supervised approach to obtain word vectors based on a training corpus. Instead of requiring dimension reduction techniques, the word vectors are dense by construction. Prediction-based representations can be divided into two categories: static representations that apply the same word vector for all appearances of a word and contextualized representations that apply a word vector dynamically based on the context it appears in.

One of the most successful static prediction-based representations is Word2Vec [52] which is an application of a continuous BoW model (cBoW) or a continuous skip-gram model. The continuous BoW model aims to predict a word based on its surrounding words. In contrast, the continuous skip-gram model aims to predict the context based on the target word. A drawback of this method is that words not contained in the training corpus are not assigned appropriate word representations. FastText [53] solves this issue by computing vectors for n-grams of characters instead of words and outperforms Word2Vec, especially when trained on an insufficient training corpus. However, the performance of static representations is limited since each word has the same vector regardless of the context.

Contextualized embeddings determine word embeddings dynamically depending on the context they appear in. ELMo [54] and Flair [55] both use bidirectional LSTMs for transforming text information into context vectors. Similar to the previous static representations, they differ in the embedding level with ELMo computing word-level embeddings and Flair computing character-level embeddings. Since both approaches rely on LSTM architecture, they are limited by the RNN drawbacks. The transformer architecture avoids this drawback. Hence, efficient state-of-the-art approaches use transformers for computing contextualized word embeddings by using the hidden representation of the encoder as word embedding. One of the most successful applications of this approach is BERT [47] which will be discussed in more detail in the following Section 2.1.3.

2.1.3 BERT

Bidirectional Encoder Representation Transformer (BERT) learns bidirectional text representations based on the encoder of its transformer architecture. Hence, it relies on multi-head self-attention layers instead of RNNs. Due to its architecture, it pre-trains deep bidirectional word representations on unlabeled text. Thus, this pre-trained model can be fine-tuned by adding an output layer for various NLP tasks such as sentiment analysis and named entity recognition.

Architecture. The architecture of BERT follows a multi-layer bidirectional Transformer encoder architecture, i.e., it stacks L identical Transformer encoder layers [45]. Each encoder contains a multi-head self-attention layer followed by a simple, fully connected feed-forward network.

Input representation. The input text is represented by WordPiece embeddings [56] instead of word embeddings which provides a balance between model flexibility and efficiency. Thus, it enables the representation of unknown words while preventing a complexity similar to character-based models.

The input is expanded by adding special tokens: a classification token [CLS] at the start of each input sequence and a separation token [SEP] at the end of each sentence. The final hidden state corresponding to the [CLS]-token is used for classification tasks as aggregated sequence representation.

The total input embedding layer consists of token, segment, and position embeddings as shown in Figure 2.5. The token embedding contains each WordPiece and special token embedding. The segment embedding is added when dealing with sentence pairs indicating which sentence a token belongs to. Lastly, the position embedding contains information on the token position within the input sequence.

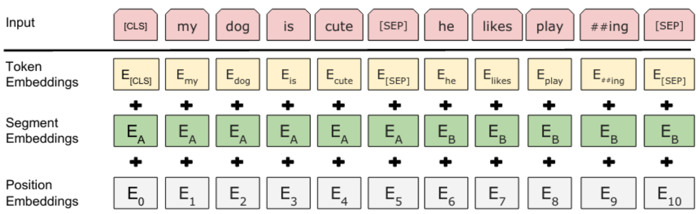


Figure 2.5: Input representation for BERT [47]

Pre-training. Using these input representations, BERT is pre-trained simultaneously on two unsupervised NLP tasks: masked language modeling (MLM) and next sentence prediction (NSP). For MLM, 15% of the input tokens are masked, and the model predicts a token based on its context tokens. Subsequently, for NSP, the model predicts if a sentence B follows an input sentence A. While the model learns the relationship of words in context with the first task, the second task enables the model to understand the relationship between sentences [47].

For English, two pre-trained models called BERT-base and BERT-large are available. Both were trained on 16 GB of data containing a corpus of 3.300 million words. While BERT-base consists of 12 encoder layers, each with 12 attention heads, BERT-large consists of 24 encoder layers, each with 16 attention heads.

Downstream tasks. In order to apply BERT to downstream tasks, two main approaches are used: feature-based and fine-tuning.

For feature-based approaches, the pre-trained model weights are used for computing BERT embeddings without updating the model weights. Similar to other text embeddings, BERT embeddings can be used as input for various models. The embeddings can be extracted using various approaches depending on the task, such as simply extracting outputs from a specific layer or taking a weighted sum of multiple hidden layers. [47]

In contrast to the feature-based approach, fine-tuning updates model weights. The model is fine-tuned on a downstream task by adding a task-specific layer on BERT's transformer stack. Fine-tuning allows the text representation to learn task-specific characteristics.

BERT-inspired models. Since the introduction of BERT, multiple variants have been developed that improve its performance for selected NLP tasks.

For instance, the Robustly Optimized BERT Pre-Training Approach (RoBERTa) [57] improves the performance for various NLP tasks, including language understanding. It pre-trains on masked language modeling only and is trained on 160GB of data. Further, domain-specific variants of BERT were introduced by pre-training on domain-specific data only to improve its performance for domain tasks. In this context, Twitter models were pre-trained to adapt to the challenging language nature of social media, such as being noisily user-generated and restricted by character limits [58]. For tasks such as semantic textual similarity (STS), BERT's architecture causes a massive computational overhead. Sentence-BERT [59] modifies BERT by using "siamese and triplet network structures for deriving semantically meaningful sentence embeddings that can be easily compared using cosine-similarity" [59]. Thus, Sentence-BERT improved the computation efficiency on STS significantly.

2.2 Topic Modeling

As a standard approach for exploratory document analysis, topic modeling is an unsupervised tool for identifying common themes and underlying narratives in text [60]. It aims to discover latent topics within the text corpus, with each describing an interpretable semantic concept [61]. Topic modeling approaches followed the evolution of NLP from conventional topic models using bag-of-words representations to neural topic models using complex prediction-based representations. The following briefly elaborates on both conventional and neural topic models.

Conventional topic models. Conventional topic models represent each document as bag-of-words and assume each document to be a mixture of latent topics. The most commonly used conventional topic model is Latent Dirichlet Allocation (LDA) [62] belonging to the Bayesian probabilistic topic models. This model generates a document using latent variables sampled from a pre-defined distribution based on Bayes' theorem. Models learn the latent variables using a Bayesian inference process. This inference process brings two major disadvantages when dealing with large data. First, it requires careful customization. Second, its complexity grows significantly with model complexity. Consequently, it does not scale on large text collections. Another type of conventional topic model is based on Non-Negative Matrix Factorization (NMF) [63] which decomposes the document-word representation into a topic-document matrix and a word-topic matrix using techniques from linear algebra and multivariate analysis.

Besides lacking scalability, the conventional methods typically rely on bag-of-words representations of documents that disregard the context of words. Consequently, these models might fail to capture the semantics of a document accurately.

Neural topic models. In order to overcome this lack of context, neural topic models have become increasingly successful in leveraging neural networks. Instead of relying on bag-of-words representations of documents, these newer topic models use contextual word and sentence embeddings, such as BERT [47], in order to capture semantic properties and embed similar texts close to each other.

Since vector embeddings of documents enable straightforward similarity measurements between documents, various topic modeling concepts are primarily based on clustering document embeddings [64, 65, 66]. Initial algorithms clustered the embeddings using the hierarchical clustering algorithm HDBSCAN [67] and then extracted the topic representations based on words close to its cluster centroid [65, 68]. Hence, this approach assumes that words close to a cluster centroid represent the cluster. However, an HDBSCAN cluster does not always lie in a sphere around the centroid.

In order to overcome this topic representation issue, neural topic modeling concepts were combined with conventional text representation techniques. For instance, BERTopic [69] computes document embeddings using Sentence-BERT[59], reduces their dimensionality using UMAP[70] and clusters them into topics using HDBSCAN [67]. In order to compute the topic representation, it merges all documents of a topic into one document and computes their TF-IDF representation (see Section 2.1.2).

2.3 Social Media Analysis

The emergence of online social media platforms has drastically changed communication between individuals and groups. Social media platforms enable different forms of communication that evolve from the platform's regulations and limitations. For instance, micro-blogging platforms limit shared information to a short-text format suitable for sharing brief information and opinions.

By analyzing social media data, researchers aim to obtain a deeper understanding of the public's behavior. Popular analysis approaches include opinion mining, information extraction, emotion analysis, and event detection. Results of these analyses can be employed for various tasks, including research in social sciences, preventing misuse of social network platforms, and supporting decision-making in companies.

Since data obtained from a social network platform differs from conventional data, applying conventional data analysis tools obtains only mediocre results [58]. Social media data is user-generated, fast-paced, and shows particularities of the respective platform. These platform-specific deviations from conventional data are a challenge for conventional analysis models. Hence, they need to be adapted to perform well on social media data, e.g., by training NLP models specifically on the platform's data.

In addition to the shared content, social media data contains meta-information on the sharing process, such as Spatio-temporal attributes, information on the content's author, and user interactions via replying, forwarding, and mentioning other users. The different forms of information give rise to various analysis approaches such as sentiment analysis on textual data, volume analysis on Spatio-temporal data, and social network analysis on interaction data. These analysis approaches can be combined for solving complex tasks such as event detection and social event analysis. In the following, these two tasks are elaborated on in more detail. Section 2.3.1 summarizes recent innovations in event detection. Following, in Section 2.3.2 approaches and results of protest analyses as example application for social event analysis are discussed.

2.3.1 Event Detection on Social Media Data

Social media platforms enable users to share information on real-world events in real-time, even before traditional news outlets release them. Detecting such events in social media data allows researchers to monitor real-world events and develop a deeper understanding of the public's perception.

Event detection on social media data can be defined as detecting an occurrence causing an abnormal change in the volume of social media data that discusses the associated topic at a specific time [71]. Due to the noisy content, fast-changing topics, and large data volumes, event detection on social media is particularly challenging [72, 73].

Based on the available information about events, event detection methods are divided into specified and unspecified methods [74]. Specified methods employ specific information on the event, such as location, time, and description. In contrast, unspecified methods employ no preliminary event information but rely on the social media stream only to detect event-related occurrences.

Specified event detection methods were introduced for detecting various types of events in offline or online manner including natural disasters [75], controversial events [76], musical concerts and festivals [77, 78]. Prior knowledge of the event, such as time and location, can be used to focus on a potentially relevant subset of tweets. By determining the characteristics of an event, event-specific features can be monitored that improve event detection. For instance, an increase in human mobility could be related to a natural disaster [79], the usage of artist and venue city names indicates musical concerts [77], and the change in acceleration in the growth of online communities could imply large-scale protests [80].

In general, offline specified event detection consists of five phases [81]. Initially, social media data is collected, and the raw data is pre-processed. Subsequently, appropriate event-specific features are selected. Based on the similarity of these features, the data is clustered into topics. Finally, an event module decides if a given topic is detected as an event. Methods and models used in the individual phases strongly depend on the events of interest.

2.3.2 Protest Analysis on Social Media Data

State-of-the-art protest analysis on social media is mainly restricted to qualitative analysis of a major online social movement related to large-scale protests such as the 'BlackLivesMatter' [13], and 'OccupyWallStreet' [82] movements in the USA, the 'umbrella' movement in Hong Kong [83], and the 'social outburst' of youths in Chile [84]. While some analyses aim to gain a deeper understanding of one specific protest [82, 13], others aim to derive transferable insights into the connection between protest, and online social movement that can be validated on other protests of similar nature [85, 16].

Prior to the analysis, social media data on the protest needs to be extracted from the respective social media platform. In general, it is collected using one or more multiple hashtags that are known to have been used in the context of the protest. For instance, hashtags containing 'BlackLivesMatter' as prefix such as 'BlackLivesMatterSTL' and 'BlackLivesMatterFerguson' were used for collecting data on the 'BlackLivesMatter' movement [13]. After collecting data, various features can be computed. Hashtags can be extracted and clustered into different types of hashtags providing information on different opinions or perspectives on the protest among the public [13]. Another popular approach constructs a social network from user interaction contained in the data. By analyzing the network structure, propositions are made on several protest aspects, including the nature of most active social media accounts [82], the decentralized structure of large-scale international social movements [22], and the network structure's effect on protest participation [85, 16]

Chapter 3

Methodology

3.1 Data Collection

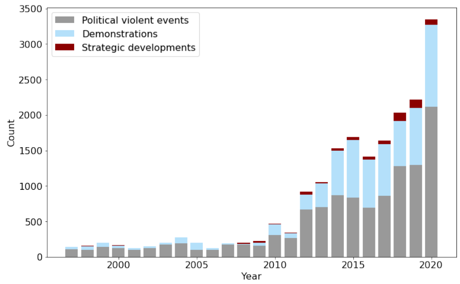
This thesis's exploratory and analytical approaches combine conflict data with social media data. For information on conflicts, political violence and protests in Nigeria a data collection by the non-profit organization "The Armed Conflict Location & Event Data Project" (ACLED)[86] was employed (see Section 3.1.1). The social media data collection was self-constructed by scraping Nigerian Twitter data (see Section 3.1.2). On top of these two main data sources, some analysis steps require context information, such as Nigerian city and state names, which was collected from various data sources (see Section 3.1.3)

3.1.1 Conflict Data

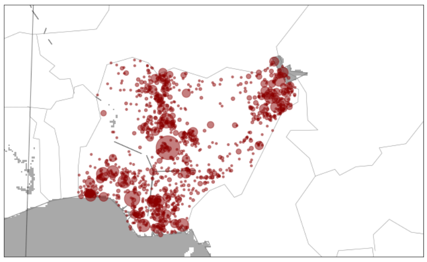
ACLED collects real-time data on political violence events (battles, remote violence, violence against civilians), demonstrations (protests, riots), and selected politically relevant non-violent events. The data is collected by data experts worldwide using local, regional, and national sources [86]. Unit of observation is the event involving actors, location, and a specific date [87].

ACLED has covered Nigeria since January 1997 [88]. The organization is constantly backcoding conflict data. Hence the number of registered events can still increase for the past years in the future [87]. Up to the extraction of the ACLED dataset on 25 October 2021, roughly 24k data points in Nigeria were registered since the start of coverage. Figure 3.1 shows the increasing number of registered events taking place per year over the past 20 years and their location distribution over Nigeria for 2020 as an example year. Besides the violent political events, the number of demonstrations increased significantly, reaching 858 protests in 2020.

This thesis will focus on the last five years of the social media era, i.e., 2016 to 2020. Due to the Twitter ban in Nigeria from 5 June 2021 to 13 January 2022, we exclude the year 2021 from consideration. In total, 2848 protests were registered in Nigeria between 2016 and 2020.



(a) Number of registered events



(b) Geographic distribution

Figure 3.1: Registered events in Nigeria by ACLED. The number of registered events in Nigeria has increased significantly since the start of coverage in 1997 up until 2020. The events are distributed over Nigeria with hotspots in certain areas such as the east-north and around the capital city Abuja in the centre. The circle sizes correspond to the number of events sharing the same location.

3.1.2 Social Media Data

The micro-blogging platform Twitter was employed as the social media data source. For this thesis, we constructed two different data collections of tweets created between 2016 and 2020: a collection of Nigerian geotagged tweets by Nigerian Twitter users and a collection of tweets restricted to selected hashtags. For scraping, the Twitter-scraping tool TWINT [89] was employed. We customized TWINT to scrape geotags and user information of the tweet authors.

Geo-restricted collection. For exploratory analysis of the Nigerian Twitter landscape, the geo-restricted collection contains only tweets by Nigerian users geotagged with a location within Nigeria. The ratio of geotagged tweets varies by topic but is assumed to be only 1-2% of all tweets available by the Twitter API [90]. Hence, this geo-restricted collection aims not to achieve completeness but rather to provide a basis for capturing local trends.

Twitter only allows a search for geotagged tweets within a radius around a central point. The radius and central point were defined so that the spanning circle covers the entire Nigeria as shown in Figure 3.2. In addition to the tweet text, further information such as geo-locations, timestamp of tweet creation, and user information of the tweet author were extracted. In order to restrict the tweet collection to Nigerian Twitter users, the scraped tweets were further filtered for users that stated a profile location containing 'Nigeria', 'NG', a Nigerian state name, or a Nigerian city name (see Section 3.1.3). This data set contains 28M of tweets in total, starting at 8.4k tweets per day on average in 2016 and reaching up to 26.8k tweets per day on average in 2020.

Hashtag-restricted collection. For a detailed analysis of selected Nigerian online social movements, tweets containing selected hashtags were scraped for pre-

Figure 3.2: Geographic circle for scraping tweets with geotags by geo-location.

determined time spans as delineated in Section 3.4. As for the geo-restricted data set, it contains the tweet text, tweet meta-data, such as the tweet timestamp and language, as well as user information on the tweet author. This data set contains 505M of tweets between 2016 and 2020 distributed over 2988 hashtags.

3.1.3 Other Data

In order to restrict the geo-referenced Twitter collection to Nigerian tweets, names of Nigerian cities and states were collected from [91]. The data set includes names of the largest 745 cities, 36 state names, and one federal capital territory.

For filtering political topics in Section 3.3, words related to 'politics' were scraped from [92]. The political word collection contains the most related 500 words and includes a similarity score for each word towards 'politics'.

For training a customized GloVe model, we employ a small corpus consisting of the first 100M characters of Wikipedia [93]. It is extended by the protest descriptions contained in the ACLED data set. The dictionary is further used for constructing a word dictionary in order to split hashtags into sub-words in Section 3.4.

3.2 Data Preparation

The tweet texts are further filtered and pre-processed using common pre-processing approaches to enable accurate text analysis.

In order to maximize information content in tweets used for exploratory analysis, only English tweets with a minimum length of ten words are considered. Further, at most 50% of the tweets' word counts are allowed to consist of emojis and user mentions. This filtering step is only applied to tasks involving NLP models. For tweet frequency analysis, all tweets are employed.

The NLP pre-processing involves converting cases to lower case, removing URLs and annotations, and translating emojis into text. An example of an original and pre-processed tweet is given below.

"I so much admire him.♡ @bukolasaraki #8thSenate
https://t.co/N0sKml1Rmv"
↓
"i so much admire him. :red heart: 8thsenate"

3.3 Topic Modeling

To understand the Nigerian Twitter landscape and its relation to conflicts and protests, we first perform an exploratory analysis on the geo-restricted Twitter data set. This exploratory approach aims to determine what Nigerians are tweeting about, which political topics are discussed online, and whether protest discussions can be identified within this extensive Twitter data set. We combine a BERT-based topic modeling approach with a political similarity rank.

The approach for topic modeling consists of three phases shown in Figure 3.3. First, a topic model is fitted on a subset of tweets sampled uniformly at random for each month. Second, the reduced topic model predicts topics for all tweets of the respective and previous month. Third, the topics are ranked based on a political similarity score. In the following, the individual steps are explained in more detail.

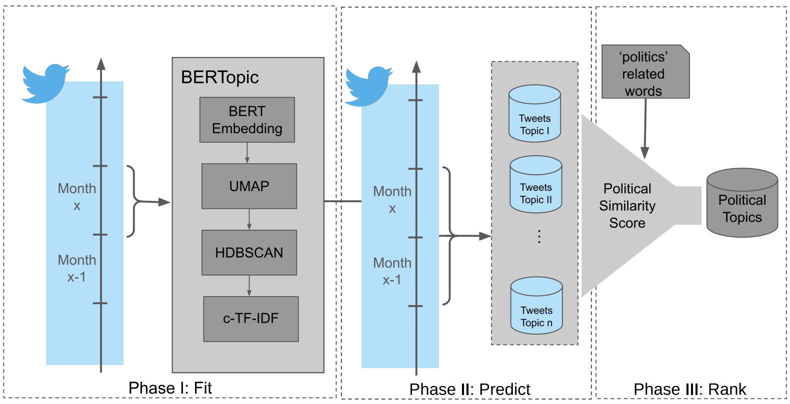


Figure 3.3: Pipeline for topic modeling consisting of three phases: Phase I fits the topic model to the tweets of a month. Phase II predicts the topics for all tweets of the month and the preceding month. Phase III ranks the topic by political similarity. Consequently, the top-ranked topics refer to the most political topics.

Fitting Topics. For modeling topics based on the geo-restricted Twitter data, the transformer-based topic model BERTopic [69] is employed. This model embeds all tweets as BERT embeddings and reduces the embedding dimension using

UMAP [70]. Based on similarities between the tweet embeddings in the lower dimension, clusters are computed via the hierarchical clustering algorithm HDBSCAN [67]. For each topic, it merges all tweets into one document and computes its topic representation via TF-IDF [50]. We model the topics on a subset of 130k tweets for each month sampled uniformly at random for efficiency.

Predict Topics. The modeled topics are used to predict topics for all tweets of each month. Topics are also predicted for the preceding month to capture the emergence of topics that started the month before. The topic for a tweet is predicted by computing its BERT-embedding, reducing the dimension using UMAP, and predicting the HDBSCAN cluster it belongs to.

Ranking Political Topics. In order to explore political discussions on Twitter, we determine to which extent the topics are related to politics based on political similarity scores. First, we extract the top 10 important words and their importance scores for each topic based on its TF-IDF representation. As topic embedding, we compute the weighted average of their customized Glove embeddings (see Section 3.1.3) using their importance scores as weights.

The scores for all topics are computed via weighted averages of the cosine-similarity between the Glove embeddings of 500 political words (see Section 3.1.3) and the topic embeddings.

Topic Analysis. For a preliminary exploration of the Nigerian Twitter landscape, we combine topics by merging the smallest topic with its most similar topic based on their TF-IDF representations until only 20 topics are left. These 20 topics per month are labeled manually with general subjects. The following general subjects are selected based on a short look-through of the data: appearance, congratulations, education, entertainment, feelings, food, online marketing, Nigeria, politics, religion, sports, traffic, Twitter, and weather.

After the preliminary exploration, we analyze what kind of political topics are discussed heavily online using the more political topics based on the originally modeled topics. We qualitatively examine the top 5 political topics containing more than 500 tweets for three sample months in 2019. For each topic, we check whether a real-world event is related to it by performing online news research based on the topic's most important TF-IDF words and its time frame.

3.4 Automatized Tweet-Protest Mapping

Results of the exploratory analysis (see Section 4.1) indicate that Twitter has been used as a communication tool by various social movements that led up to offline collective actions in Nigeria.

We introduce an approach for automatically detecting and mapping online social movements related to protests of interest contained in the ACLED data set in order to enable a quantitative comparative analysis.

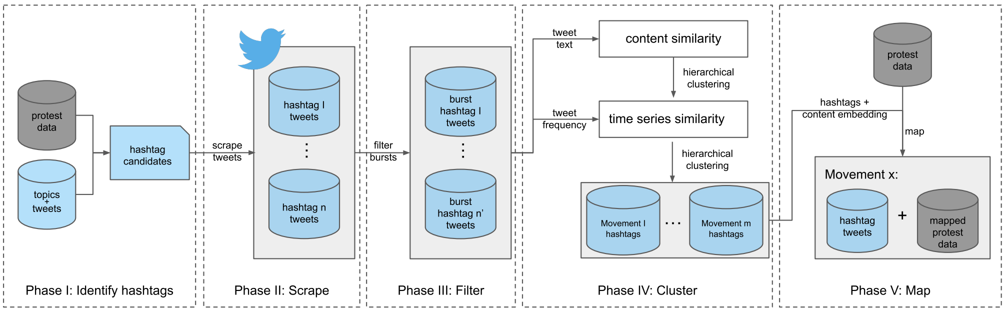


Figure 3.4: Process of Automatized Tweet-Protest Mapping: Potentially relevant hashtags are identified using a comparison of protest information and modeled topic representations. Tweets for the relevant hashtags are scraped, filtered, and clustered into online social movements that get subsequently mapped to protests.

The approach consists of five sequential steps shown in Figure 3.4. At first, hashtags that are potentially related to protests are identified. The corresponding tweets are extracted, followed by filtering and clustering the hashtag-tweet collections into online movements. Finally, these online movements are mapped to protests. In the following, the individual steps are described in more detail.

I: Identifying Hashtag Candidates. For determining hashtags that are potentially related to protests, topic models computed on the geo-referenced tweet collection in Section 3.3 are connected to the protest data set via content comparison.

For each protest, the protest data set contains a short note by the registrant that describes its cause and the protesters, as well as police intervention if applicable. We embed the protest note by computing the average GloVe embedding of all words contained in the protest note, excluding stopwords, most common words, and temporal information. For each topic, we extract the top 10 important TF-IDF words and embed the topic by computing a weighted average of these words using the word importance scores as weights.

By computing the cosine-similarity between the protest and topic embeddings per month, we determine the topics in the time interval of the protest that are similar to the protest descriptions. All hashtags occurring in the five most similar topics per protest are collected in a monthly hashtag candidate set. Hence, the monthly hashtag candidate set contains hashtags that are potentially related to protests in the respective month.

II: Scraping Tweets for Hashtag Candidates. For each hashtag in the candidate sets, a time span is determined by identifying the first and the last month it occurs in the candidate sets. By considering the months of occurrences and the months in between, we take long-term online social movements into account that lead to various timely separated protests.

Since the geo-referenced data set is strongly restricted by the geo-tag requirement, it lacks complete coverage of online social movements. Here, we omit this requirement since we assume that hashtags related to Nigerian protests are specific enough to cover Nigerian events only. Hence, we can enrich the social media data set by scraping all tweets containing the hashtag during the determined time span for each hashtag. The resulting hashtag-restricted data set is described in more detail in Section 3.1.2.

III: Filtering Hashtag Candidates. For omitting non-bursting hashtags from the following steps, we apply Kleinberg's burst detection algorithm [94]. The algorithm identifies time periods in which a target event, in this case, tweet posting, is uncharacteristically frequent.

IV: Clustering Hashtag Candidates to Movements. Online social movements often use not only one hashtag but multiple hashtags containing either different information on the movement's cause or slight variations of the same hashtag. Thus, instead of individual hashtags, online movements are the unit of interest for mapping tweets to protests. Hence, we cluster hashtags into online movements based on content and time series similarity. For measuring content similarity, we compute TF-IDF representations on the hashtag-restricted data set, which determines for each word how often it appears in a hashtag-tweet collection relative to all hashtag-tweet collections. We compute the similarity between these TF-IDF representations and apply the hierarchical clustering algorithm HDBSCAN [67] in order to determine clusters of hashtags related to similar content. Since some content is repeatedly associated with protests such as sexual harassment or police violence, these clusters of hashtags get further divided using time series similarity. Hashtags belonging to the same online social movement trend roughly simultaneously. Thus, for all hashtag-tweet collections in a content cluster, we compute the euclidean distances between their normalized daily tweet counts. On this distance measure, we apply hierarchical clustering once more to identify clusters with roughly simultaneous trends. Finally, we obtain clusters of hashtags and tweets that discuss similar content and show similar temporal trends. In the following, we will refer to these clusters as movements.

V: Mapping Movements to Protests. An online social movement can lead to one or multiple protests, while one protest is typically related to at most one online social movement. Thus, we map each protest to at most one movement. For this, we compare two different approaches: string and embedding comparison.

The string comparison approach assumes that the set of hashtags or, respectively, the set of the best describing words of an online movement related to a protest contains the most important buzzwords that characterize the protest, such as its main trigger, its goal, or the main affected. We map the protest to the online movement with the highest ratio of best-describing words occurring in the protest note. A hashtag occurs in the protest note if at least half of its subwords, excluding stopwords and most common hashtag subwords, are contained in the protest note. We compute hashtag subwords by probabilistically splitting hashtags based on relative word frequencies. The most likely split of a hashtag maximizes the aggregated probability of the individual words and is thus easily computable using dynamic programming. As a dictionary, we employ an English Wikipedia dictionary that we extend with words contained in the protest notes in order to cover protest-specific and Nigerian words and names (see Section 3.1.3).

The embedding comparison approach assumes that the aggregated embedding of a movement's most important TF-IDF words is similar to the embedding of the words describing the related protest. Similar to the topic embeddings, we compute the movement embeddings by computing the weighted average of their most important TF-IDF words. For each movement, we extract the best describing words by applying a minimum threshold of 0.01 to its TF-IDF word importance scores. Table 3.1 shows the 20 most important words and their importance score for one example movement. The protest note is embedded

TF-IDF Word	Importance Score
onnoghen	0.125
cjn	0.101
cct	0.029
suspension	0.020
onnoghens	0.017
justice	0.015
njc	0.013
walter	0.013
court	0.012
onnoghensuspension	0.011
chief	0.011
onnoghentrial	0.010
suspended	0.009
trial	0.009
buhari	0.009
law	0.009
mbuhari	0.008
nigeria	0.007
tyrantbuhari	0.007
judiciary	0.007

Table 3.1: Most important words for sample online movement against the suspension of the chief of justice Walter Onnoghen using the following hashtags: '#onnoghensuspension', '#onnoghen', '#cjn'. Only words with importance scores above 0.01 are considered during the embedding comparison.

by determining the relevant words excluding stop words, most common protest words, and temporal information. We compare two techniques for computing the similarity between protest words and movement embeddings: For the first technique, we map the protest to the most similar movement based on cosine-similarity between the mean of the protest word embeddings and the movement embeddings. For the second technique, we compute the cosine-similarity between the movement embeddings and each protest word separately, combine the mean-centered similarities per movement via max-pooling and map the protest to the most similar movement. In other words, the first approach maps the protest to the movement that is most similar to the combination of protest words. In contrast, the second approach maps the protest to the movement that is highly similar to a protest word that is not similar to most other movement embeddings.

The mapping approaches are compared on a subset of 200 protests sampled uniformly at random from the ACLED protest data set of 2019. We manually labeled protests with hashtags, if applicable, by performing a Twitter search for protest words during the days around the protest. The mapping quality is determined by computing recall, precision, and f1-score of the mappings.

3.5 Inter-Protest-Analysis

To identify similarities and dissimilarities among the movement-protest pairs, we perform a comparative analysis of selected features. We aggregate features per day to avoid measuring daily fluctuations. Further, we refer to the number of tweets per day as Twitter volume.

Preliminary Pattern Analysis.

Initially, we identify categories of movement-protest mappings based on movement duration and the temporal relation between the Twitter volume peak date d_{peak} and the protest $d_{protest}$.

First, we analyze the distribution of movement duration among all movement-protest mappings. Here, the movement duration is defined as the number of days between the start of the initial burst detected by Kleinberg's burst algorithm [94] (see 3.4) and the end of the last detected burst.

Further, we analyze the Twitter volume trend and the distribution of movement-protest pairs among three categories of relationships between Twitter volume peak date d_{peak} and the respective protest date $d_{protest}$ as shown in Table 3.2. The date d_{peak} of the Twitter volume peak is determined by identifying the date of maximal Twitter volume within ten days before and after the protest date $d_{protest}$. Restricting this time interval enables the detection of multiple peaks of a movement corresponding to timely separated protests. The number of days between the Twitter volume peak date and the protest date is referred to as $\Delta = d_{peak} - d_{protest}$.

Category	Condition
I	$d_{peak} < d_{protest}, \Delta > 0$
II	$d_{peak} = d_{protest}, \Delta = 0$
III	$d_{peak} > d_{protest}, \Delta < 0$

Table 3.2: Categories of short-term movement-protest mappings based on the temporal relation between Twitter volume peak date d_{peak} and the respective protest date $d_{protest}$.

Category I refers to movements that emerge before the protest takes place. Thus, they enable a retracing of the online discussion leading up to the protest. Movement-protest pairs of Category II simultaneously emerge, i.e., the online movement peaks on the day of the protest. Lastly, Category III refers to online movements emerging after a protest.

Feature Computation.

For the comparative analysis, we compute various features employing different aspects of the collected data. We capture Twitter volume trends, user engagement, and the discussion mood.

Twitter Volume Trend. The general relationship between movement trend and protest day is analyzed via the Twitter volume trend. For each movement-protest pair, we determine the Twitter volume at its peak day, referred to as n_{peak} and on the protest day, referred to as $n_{protest}$.

User Engagement. We examine the user engagement via multiple features using the user metadata and the @mention meta-data of tweets.

Based on the user meta-data, we compute user activity, the ratio of new movement participants per day, user accounts' ages, and the ratio of Nigerian users. First, for each day of a movement, the user activity is determined, i.e., the number of tweets a user creates per day. Further, for analyzing the stability of the user base, the ratio of users joining the movement's online discussion for the first time is computed. In order to examine the age of Twitter accounts and determine the number of "freshly" created accounts, we compute the age of the user account on the protest date. Here, we define "freshly" created accounts as accounts created within ten days preceding the protest. Further, for analyzing the extent of national vs. international attention of a movement, we determine the ratio of users that state user locations assignable to locations within Nigeria. Since the user manually enters the user location in clear text, we renounce interpreting the total numbers or ratios but focus only on differences between them during comparisons.

Using the @mention-metadata of tweets, we analyze the interaction among movement participants. We construct movement networks from the set of annotations per day. Twitter accounts are modeled as nodes connected via a directed edge if a Twitter account mentions another. Hence, the resulting network is a directed multigraph. We analyze the network's hierarchy structure by determining the extent to which the networks exhibit a flow hierarchy structure [95]. The hierarchy metric takes on values in $[0, 1]$ with 1 referring to a strong hierarchical structure and 0 referring to the absence of a hierarchical structure. Further, we analyze the centrality of networks using two centrality metrics: indegree and outdegree centrality. Indegree centrality measures how central a node is regarding the number of incoming edges. In contrast, outdegree centrality measures the centrality of nodes regarding the number of outgoing edges. Consequently, nodes with a high indegree centrality correspond to Twitter accounts being frequently mentioned by others, while nodes with a high outdegree centrality correspond to Twitter accounts mentioning others frequently. For network visualizations, we use the analysis and visualization software *visone* [96] and compute layouts via stress minimization.

Mood of Discussion. In order to grasp the mood of a movement discussion, the tweet sentiments are computed. For this, we employ a RoBERTa-based sentiment classification model trained on Twitter data only [58]. The classification model assigns each tweet a probability score for being of negative, neutral, or positive sentiment.

Feature Analysis.

We analyze features by aggregating over movement-protest pairs. Based on the aggregated features, we compare trends as the movements evolve and analyze dependencies on the number of days Δ between features of the Twitter volume peak day and the protest day. Further, we analyze the development of selected aggregated features over the years. Lastly, we compare aggregated features between two selected categories of movement-protest mappings.

In addition, we analyze the nature of protests mapped to online movements by comparing their geographic distributions and their distribution among protest types (peaceful, with intervention, excessive force) to the respective distributions of all registered protests.

Chapter 4
Results

This chapter presents the results in three sections. First, the modeled topics are explored in Section 4.1. Subsequently, the mapping approaches of tweets to protests are evaluated in Section 4.2. Finally, we provide the results of the comparative analysis in Section 4.3.

4.1 Topic Modeling

In general, our approach models a large number of small sized topics. The number of modelled topics increases from 850 topics on average per month in 2016 to 1050 topics on average per month in 2020. Similarly, the average number of tweets per topic increases from 45 tweets per topic in 2016 to 65 tweets per topic in 2020.

The manually determined distribution of categories among the reduced topics for 2019 as sample year is shown in Figure 4.1. Predominantly, entertainment topics are discussed, followed by feelings and sports. In addition, politics, appearance and religion are frequent topics on Nigerian Twitter.

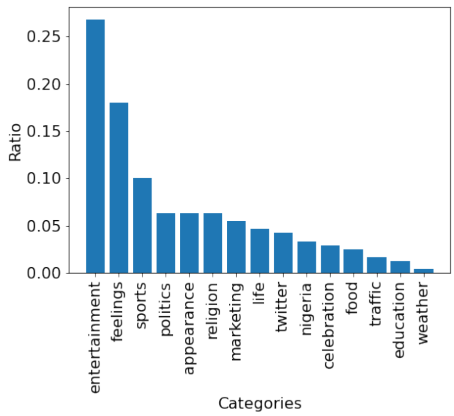


Figure 4.1: General topic categories in 2019 identified by manually labeling the 20 reduced topics with selected categories.

Table 4.1 presents the 5 most political topics containing more than 500 tweets for three months of 2019. Each topic description is formulated based on a qualitative news research using the topic's most important TF-IDF words and the topic's time frame.

2019-01	2019-02	2019-03
Nigerian election	Nigerian government	International women's day
Presidential debate	American politics	Governor elections
Election campaigns	PDP campaign	Lagos building collapse
Buhari	Vote motivation	Onnoghen tribunal
Onnoghen suspension	Nigerian elections	Kano elections

Table 4.1: Mapping of top 20 ranked topics in regards of political similarity to real-world events for three example months in 2019.

Notably, for each month, the most political topics in Table 4.1 contain at least one topic related to a protest. For instance, a protest against the suspension of the Chief of Justice Walter Onnoghen took place in January 2019, multiple protests encouraging peaceful voting in the Nigerian elections were organized in February 2019, and protests against irregularities during the governor elections took place in March 2019. However, for none of these manually detected topics related to protests the word 'protest' appears in the 10 most important topic words as shown in Table 4.2.

	Onnoghen suspension	Vote in Nigerian Elections	State elections
1	cjn	vote	apc
2	justice	nigeriadecides2019	pdp
3	court	elections	parties
4	onnoghen	nigeria	candidate
5	njc	electionday	states
6	suspension	votenotfight	inconclusive
7	cct	civic	incumbent
8	law	wisely	winning
9	suspended	decide	adamawa
10	walter	polls	elections

Table 4.2: Top 10 words for three example topics of three different months in 2019 that are related to protests.

4.2 Automatized Tweet-Protest Mapping

Results of the mapping quality assessment are summarized in Table 4.3. The highest metric scores (recall of 0.655, precision of 0.593, f1 of 0.623) are achieved by the string comparison approach using only the movements' hashtags. The best performing embedding comparison approach (recall of 0.600, precision of 0.375, f1 of 0.416) uses GloVe Embeddings and computes the similarity per protest word before max-pooling the similarities. In contrast, the embedding comparison approach using GloVe Embeddings but computing the similarity using the mean protest word embedding obtains the lowest scores for all three score measurements (recall of 0.050, precision of 0.25, f1 of 0.084).

Approach	Recall	Precision	F1
SC Hashtag	0.655	0.593	0.623
EC GloVe words	0.600	0.375	0.461
SC Content words (>4 words)	0.625	0.313	0.417
EC Bert sentence	0.294	0.313	0.303
SC Content words (>3 words)	0.250	0.313	0.278
EC GloVe mean	0.050	0.250	0.084

Table 4.3: Scores of mapping approaches: String comparison approaches (SC) outperform embedding comparison approaches (EC). Best scores are achieved by the string comparison approach using only words occurring in hashtags.

In general, the results show that for the best three performing approaches the recall is significantly higher than the precision. Further, simple string comparison approaches perform better than embedding comparison approaches.

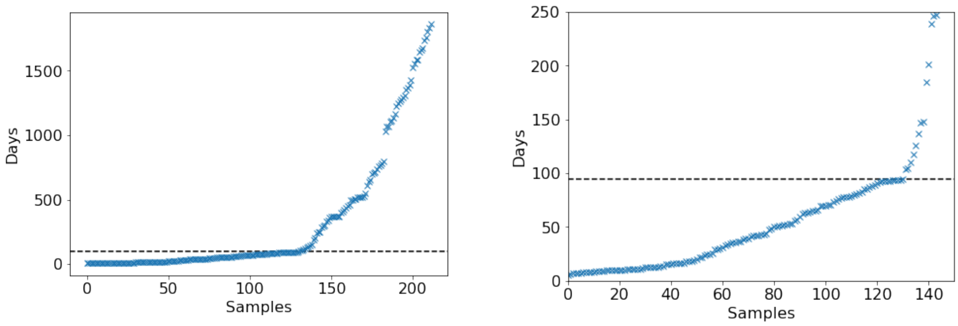
4.3 Inter-Protest-Analysis

In this section, we present the results of the comparative analysis of movement-protest mappings. First, we examine categories of mappings showing different general characteristics in Section 4.3.1. In Section 4.3.2, the results of the detailed analysis of one category are presented. Further, in Section 4.3.3, we explore the dissimilarities between two categories. Last, we examine the nature of all mapped protests in Section 4.3.4.

4.3.1 Categories of Movement-Protest Mappings

Duration Analysis. Results of the online movement duration analysis suggest a division of movements by duration into short-term and long-term movements. The duration trend in Figure 4.2 shows a significant corner at the duration of around 95 days. Hence, we define short-term movements as movements lasting at most 95 days whilst long-term movements last longer than 95 days.

Figure 4.3 shows example Twitter volume trends for short-term (Figure 4.3a - Figure 4.3c) and long-term movements (Figure 4.3d). Whilst short-term movements are typically related to to one or only a couple of temporally close protests, long-term movements, including the prominent '#endsars', '#bring-



(a) Duration of Movements

(b) Zoomed Duration of Movements

Figure 4.2: Duration of Movements sorted by number of days. The duration increases slowly until reaching a duration of 95 days. Above this threshold, the movement's duration varies from multiple months to multiple years.

backourgirls' and '#revolutionnow' movements, are related to multiple timely separated protests. Short-term movements accounting for 61% of mapped online movements are more frequent than long-term movements accounting for 39%.

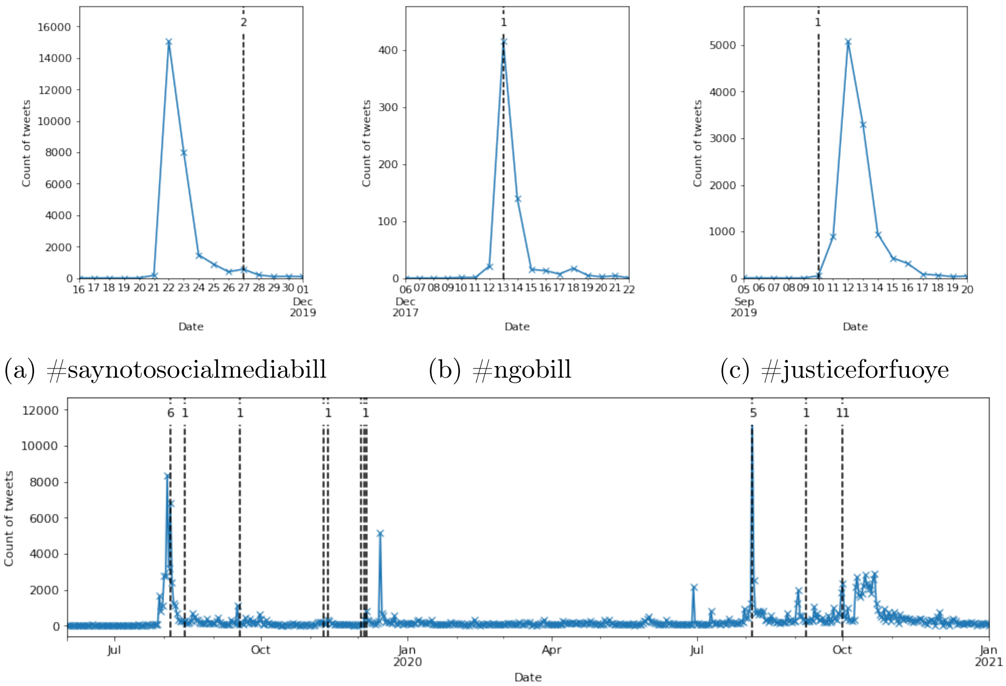
Category Analysis. Since long-term movements typically include multiple timely separated protests, the focus of category analysis lies on short-term movements. We refer to the categories described in Table 3.2 as short-term I (ST_I), short-term II (ST_II) and short-term III (ST_III). Figures 4.3a - 4.3c show sample Twitter volume behaviors for each category.

The distribution of categories among the movement-protest pairs is shown in Table 4.4. Category ST_I is most frequent accounting for 63.2% of all movement-protest pairs whilst the least frequent category is ST_II accounting only for 11.5% of the movement-protest pairs.

Category	Condition	Ratio
ST_I	$d_{peak} < d_{protest}$	63.2%
ST_II	$d_{peak} = d_{protest}$	11.5%
ST_III	$d_{peak} > d_{protest}$	25.3%

Table 4.4: Distribution of Categories among short-term movement-protest pairs. Category ST_I ($d_{peak} < d_{protest}$) is most frequent accounting for more than 63% of movement-protest pairs.

Since ST_I mappings consist of movements that emerged already before a protest takes place, they enables us to retrace the online behavior leading up to a protest. Consequently, we focus on ST_I mappings in the following detailed analysis of movement-protest mappings.



(a) #saynotosocialmediabill (b) #ngobill (c) #justiceforfuoye

(d) #revolutionnow

Figure 4.3: Sample Twitter trends for short-term movements (a - ST_I, b - ST_II, c - ST_III) and long-term movements (d). The Twitter volume is plotted as blue line, while the mapped protests are plotted with vertical black dashed lines. The number of protests per date is annotated at the top of the protest line.

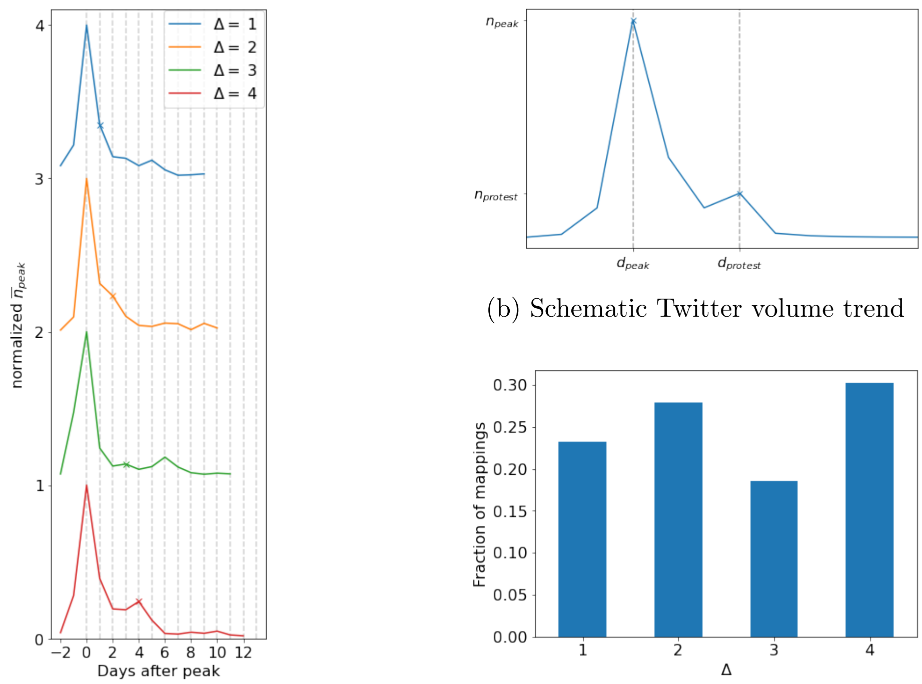
4.3.2 Characteristics of Short-Term Movements

In this section, we present the results of analysing characteristics of ST_I movement-protest mappings ($d_{peak} < d_{protest}$). Results include insights into common Twitter volume trends, user engagement behavior and the mood of the online discussion.

Twitter Volume Trend

Figure 4.4a shows the average Twitter volume trend of movement-protest mappings for different values of $\Delta = d_{protest} - d_{peak}$. In general, the movement emerges rapidly until reaching a peak after two or three days. Subsequently, it follows roughly a power-law decay which is interrupted by a short and lower second peak on the protest day. Based on these observations, we establish a schematic Twitter volume behavior shown in Figure 4.4b. The Δ-values are roughly equally distributed among the movement-protest mappings as shown in Figure 4.4c.

Dependencies between the number of tweets on the Twitter volume peak and the protest day and the number of days Δ in between are presented in Figure 4.5. First, Figure 4.5a shows that with an increasing number of tweets on the



(a) Average Twitter volume trend

(b) Schematic Twitter volume trend

(c) Δ-Distribution of mappings

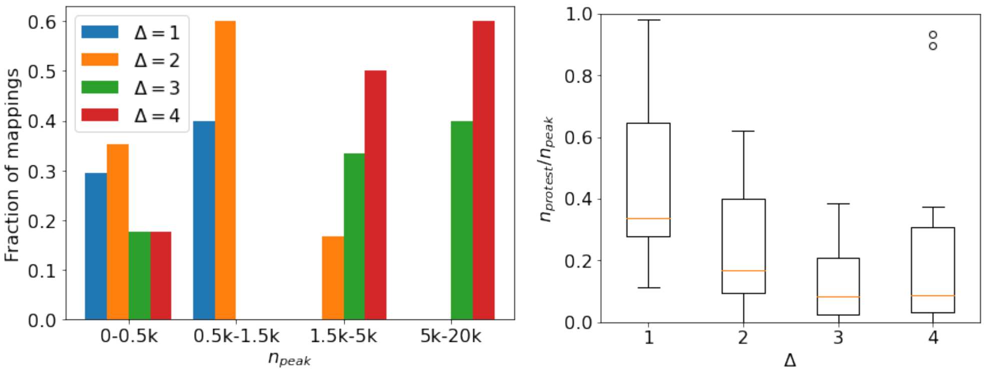
Figure 4.4: General characteristics of Type I mappings. The average Twitter volume trend for various values of Δ (Figure a) indicates a second peak in Twitter volume on the protest date. This behavior is schematized in Figure b. Figure c shows the distribution of Δ-values among the movement-protest pairs.

peak day the fraction of mappings with more days between peak and protest day increases. Further, Figure 4.5b illustrates that the ratio between tweets on the protest day and tweets on the peak day decreases with increasing Δ, i.e. the second Twitter volume peak is lower in relation to the first peak when the temporal distance is higher.

Hence, if a mapped online movement emerges to a high Twitter volume peak (>1.5k tweets), the protest is likely to happen multiple days after the peak. Further, the Twitter volume on the protest day is likely to reach only a small fraction of the Twitter volume peak. In contrast, if a mapped online movement emerges to a low Twitter volume peak (<1.5k tweets), the protest is likely to happen within the next two days. The Twitter volume decreases but reaches a larger fraction of the Twitter volume peak compared to the previous setting.

User Engagement

User engagement is analysed via user activity, the fraction of new discussion members and the interaction between users based on annotations in tweets.



(a) n_{peak}-Δ dependency

(b) Δ- $n_{protest}/n_{peak}$ dependency

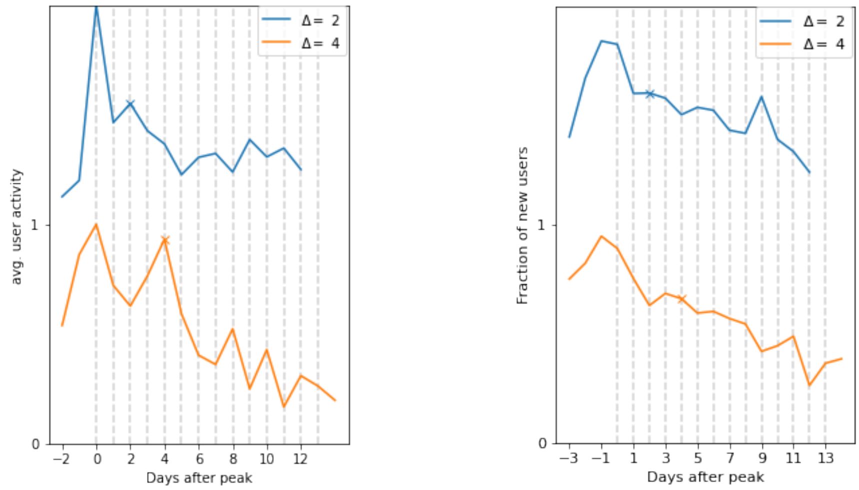
Figure 4.5: Dependencies between the number of days Δ between peak and protest date, the Twitter volume n_{peak} at its peak and the Twitter volume $n_{protest}$ on the protest date. Figure a) indicates that with increasing Twitter volume peak the number of days between peak and protest date increases. Figure b) shows a decreasing ratio between Twitter volumes on the protest and peak day with increasing number of days between them.

User Activity. Average normalized trends of the daily user activity, i.e. the number of tweets per user per day, are shown in Figure 4.6a for two temporal distances (Δ) between Twitter volume peak and protest day. In the time period between Twitter volume peak and protest day the average trend exhibits an increased user activity. After the protest it declines again. On both, Twitter volume peak and protest day, the user activity peaks with a mean value of 1.64 ± 0.25 tweets per day on the Twitter volume peak day and 1.33 ± 0.15 tweets per day on the protest day. Hence, on average, users are more active during the initial emergence of the movement than on the protest day.

New Participants in Discussion. Figure 4.6b shows the fraction of users joining the discussion for two temporal distances (Δ) between peak and protest day. It reaches a peak during the initial emergence of the movement typically a day before the Twitter volume peak day. Afterwards, it slowly decreases which is interrupted by a slight increase the day before or on the day of the protest itself.

User Interaction. In the following, we present the results of analysing the @mention-interaction on the Twitter volume peak and protest day. Interaction networks constructed based on annotations belonging to one sample online movement are shown in Figure 4.7 for both, the Twitter volume peak day and the protest day.

Aside from the difference in network size, both networks exhibit a centralized hierarchical structure. Whilst the interaction network of the example in Figure 4.7a has only one node at the hierarchical top, the interaction network of the example in Figure 4.7b has multiple nodes at the top of the hierarchy resulting



(a) User activity (b) Fraction of new discussion members

Figure 4.6: Average trends of user engagement for two example values of Δ. The user activity is shown in Figure a) and indicates peak in user activity on the Twitter volume peak and protest day. The fraction of new users joining the movement's discussion is plotted in Figure b) and shows a peak on the day before the Twitter volume peak day as well as a slight peak the day before the protest.

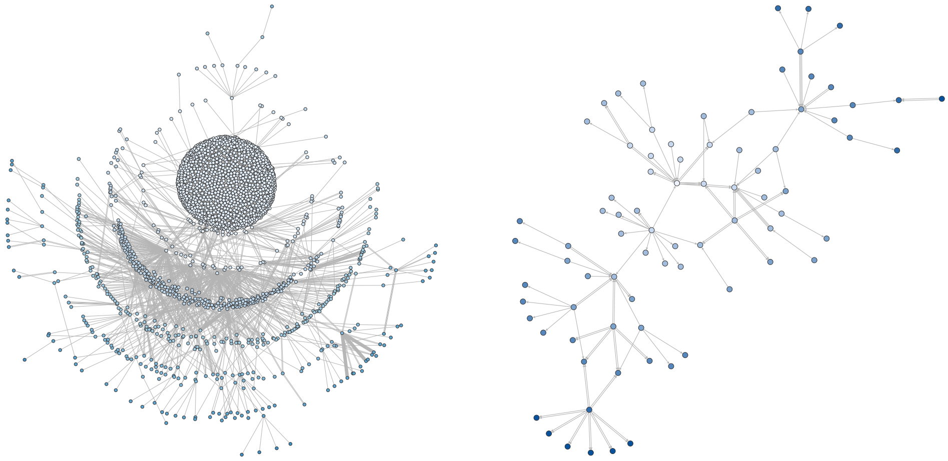
in a broader hierarchical structure. Further, each node that is high in hierarchy is surrounded by a group of nodes mentioning only this node.

The median ratio of interaction to Twitter volume is similar for both Twitter volume peak and protest day with a ratio of 0.287 ± 0.054 on the former and 0.274 ± 0.14 on the latter. Hence, the interaction volume behaves roughly similar to the Twitter volume.

The extent of the hierarchical network structure is measured as flow hierarchy. Across all networks the average flow hierarchy is slightly higher on the Twitter volume peak day with 0.978 ± 0.013 than on the protest day with 0.970 ± 0.013. Nonetheless, the flow hierarchy values are high on both days and thus, all networks exhibit a strong hierarchical structure.

Across all networks, high indegree centrality values are focused on a few prominent nodes, each of them related to a verified Twitter account. Table 4.5 shows the ten nodes for both peak and protest day that most frequently exhibit a high indegree centrality value across all networks.

Both node lists are lead by the Twitter account 'mbuhari' belonging to the current Nigerian president. On the Twitter volume peak day this account belongs to the top 10 most frequently mentioned nodes for 53% of the mapped movements. Further, both centrality lists contain the Twitter accounts of the the two largest Nigerian parties and the Nigerian police. Whilst the node list on the Twitter volume peak day contains predominantly accounts of politically active people



(a) #saynotosocialmedia - peak day (b) #saynotosocialmedia - protest day

Figure 4.7: Interaction networks on Twitter volume peak day in Figure a) and on protest day in Figure b). Whilst the interaction is stronger on the Twitter volume peak day, both networks exhibit a distinct hierarchical structure of interaction.

such as top politicians and activists, the node list on the protest day contains accounts of conventional news channels and musicians.

Name	Entity
mbuhari	president
ngrpresident	president
gidi_traffic	social media news
policeng	state organ
profosinbajo	vice president
officialpdpnig	party
youtube	social media
apcnigeria	party
segalink	human right activist
atiku	former vice president

(a) Twitter volume peak day

Name	Entity
mbuhari	president
policeng	police
ngrpresident	president
officialpdpnig	party
cnn	news
youtube	social media
davido	musician
apcnigeria	party
realdonaldtrump	foreign politican
falzthebahdguy	musician

(b) Protest day

Table 4.5: Nodes most central (indegree) most frequently across all networks. The account of Muhammadu Buhari, the Nigerian president, leads both lists.

In contrast to the indegree centrality, the outdegree centrality is more movement specific with rarely one node being central across multiple networks. The nodes with high outdegree centrality are typically related to personal accounts of conventional Twitter users that are not of public interest, i.e. they are not verified accounts.

Mood of Twitter Discussion

The results of sentiment analysis are shown in Table 4.6. Across all movment-protest pairs, the sentiment scores are roughly stable over the relevant time period of the movement between Twitter volume peak and protest day with only small deviations. The average sentiment score is highest for the neutral sentiment class, followed by the negative sentiment class. The positive sentiment score is significantly lower. On the protest day the average sentiment scores only slightly deviates from the average sentiment scores on the Twitter volume peak day. Both, the negative and positive sentiment scores slightly decrease whilst the neutral sentiment score slightly increases.

Sentiment	$\overline{\text{sent}}$	$\overline{\text{sent}}_{protest} - \overline{\text{sent}}_{peak}$
negative	0.349 ± 0.064	-0.030
neutral	0.496 ± 0.056	+0.031
positive	0.155 ± 0.048	-0.001

Table 4.6: Average sentiment scores and their change between Twitter volume peak day and protest day. Neutral and negative sentiment are prevailing. Between Twitter volume peak and protest day the sentiment score changes only slightly towards more neutrality.

Consequently, online discussions mapped to protests are of neutral-negative nature and show only slight variations between the Twitter volume peak day and the protest day.

Development over the years

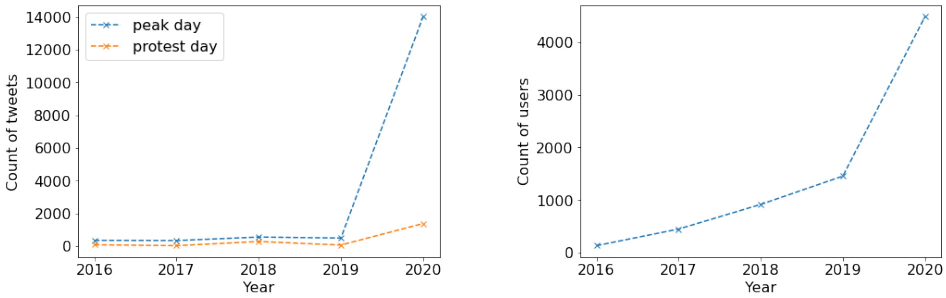
Lastly, we present the results of analysing the temporal development of features. Here, we focus on the Twitter volume and the age of the Twitter account.

Figure 4.8a shows that the average Twitter volume on its peak day grows over the years with a drastic increase in 2020. Similarly, in Figure 4.8b, the average number of Twitter users created within the ten days preceding the protest grows and shows also a significant increase in 2020. Consequently, both characteristics indicate a drastic change of the nature of online social movements in 2020.

4.3.3 Short-Term Sub-Movements in Long-Term Movements

Pattern comparison of long-term movements and short-term movements indicates that long-term movements contain multiple short-term sub-movements for each timely separated group of protests. Figure 4.9 illustrates one short-term sub-movement within an example long-term movement.

In the following, we present characteristics that show different behavior for Category I short-term movements (ST_I) and Category I short-term sub-movements of long-term movements referred to as LT_I movements.



(a) Twitter volume peak trend

(b) 'Freshly' created Twitter users

Figure 4.8: Development of Twitter volume peak and fraction of 'freshly' created users over the considered years. Both properties show an increase in 2020 which indicates a change in Twitter usage in Nigeria.

The average user activity is shown in Table 4.7a. Whilst the average user activity for ST_I movements only slightly decreases from Twitter volume peak to protest day, the average user activity for LT_I movements drops drastically. Further, LT_I user activity is notably higher for both days whilst the standard deviation is also significantly higher.

In addition, the average fraction of users joining the movement is lower for LT_I movements than for ST_I movements on both peak and protest day as shown in Table 4.7b.

	Peak Day	Protest Day
ST_I	1.64 ± 0.25	1.33 ± 0.15
LT_I	2.51 ± 2.11	1.78 ± 0.91

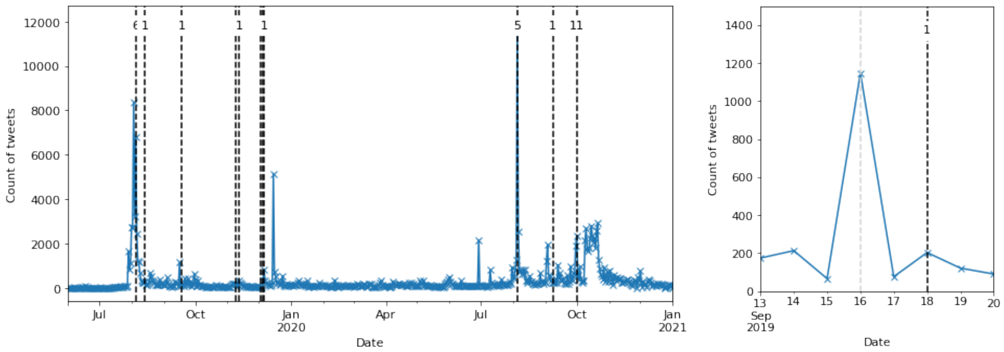
(a) User activity

	Peak Day	Protest Day
ST_I	0.88 ± 0.11	0.61 ± 0.18
LT_I	0.66 ± 0.28	0.48 ± 0.25

(b) Fraction of new discussion participants

Table 4.7: Comparison of user engagement for short-term movements of category I ($d_{peak} < d_{protest}$, ST_I) and short-term sub-movements of category I ($d_{peak} < d_{protest}$, LT_I). User engagement is significantly higher for short-term sub-movements of long-term movements (LT_I), whilst the user engagement for short-term movements (ST_I) exhibits smaller standard deviation.

Lastly, the ratio of Twitter users with user locations in Nigeria is lower for LT_I movements than for ST_I movements by roughly 15%. We renounce to state total ratio numbers since the user location is highly noisy due to the manual clear text entries. However, since we apply the same approach for both LT_I and ST_I we can interpret the difference between the two ratios.



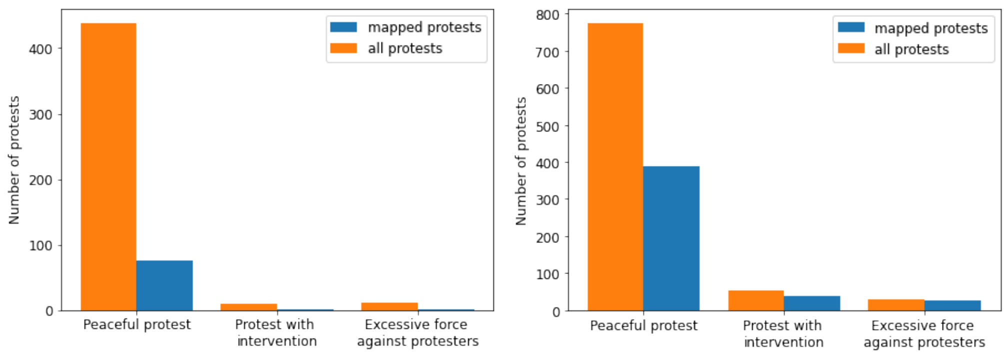
(a) Long-term movement (b) Movement segment

Figure 4.9: Short-term sub-movements in long-term movements for the example movement #revolutionnow. Figure a) shows the movement's entire Twitter volume trend (blue lines) with mapped protests (black dashed vertical lines). Figure b) shows a short-term sub-movement of the full movement.

4.3.4 Analysis of Mapped Protests

In this section, we present the results of analysing the nature of protests that were mapped by our approach.

The distributions among the type of protests for 2016 and 2020 are shown in Figure 4.10. In general, the total number of all protests increased between 2016

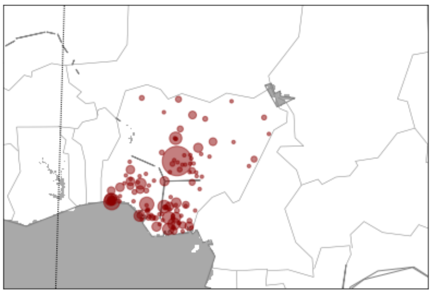


(a) 2016 (b) 2020

Figure 4.10: Distribution of protest types among all protests and only the mapped protests in 2016 and 2020. The capturing of non-peaceful protests improves from 2016 to 2020 making them the most reliably mapped protest in 2020.

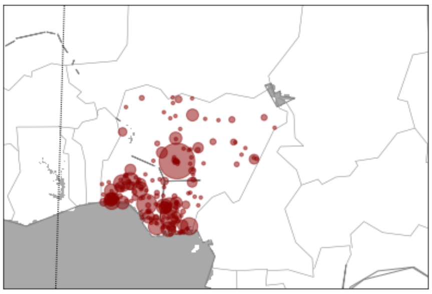
to 2020 whilst the fraction of peaceful protests decreased. For 2016, our approach is able to map only a small fraction (17%) of the protests and captures mainly peaceful protests. In contrast, for 2020, our approach maps 52% of the protests and most reliably captures non-peaceful protests.

Figure 4.11 shows the geographic distribution for the first considered year 2016 and the last considered year 2020 for all registered protests as well as for only the mapped protests. For both years, the largest number of all registered and all mapped protests take place in the capital city Abuja. Whilst only a few protests outside of the capital city are mapped for 2016, the geographic distribution is significantly broader for 2020. Especially a large number of protests on the south coast of Nigeria are captured by our mapping approach in 2020.

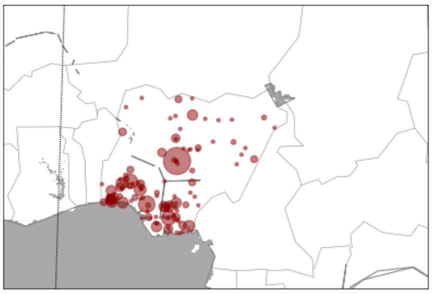


(a) 2016 - all protests

(b) 2016 - mapped protests



(c) 2020 - all protests



(d) 2020 - mapped protests

Figure 4.11: Geographic Distribution of Protests for 2016 and 2020: Figures a) and c) show the geographic distribution of all Nigerian protests contained in the ACLED data set whilst Figures b) and d) show the protests mapped by our automatic mapping pipeline. Whilst the mapped protests focus mainly on the capital city Abuja in 2016, the mapping pipeline is able to additionally capture protests in the south of Nigeria in 2020.

Chapter 5

Discussion

In this chapter, we discuss the results presented in Section 4, their limitations as well as potential directions for future work. We discuss the results of topic modeling in Section 5.1. Subsequently, we interpret the results of the mapping approaches comparison in Section 5.2. In Section 5.3, we discuss and contextualize the results of the comparative analysis across movement-protest mappings. Finally, we present the limitations of our approaches and discuss potential directions for future work.

5.1 Topic Modeling

A broad range of topics is discussed on Twitter in Nigeria. The distribution of categories among the reduced topics indicates that Nigerians use Twitter mainly as a tool for discussing entertainment-related topics such as music, TV shows, and radio shows. Additionally, they frequently communicate their feelings and discuss sports events. However, also political topics account for a non-negligible share of topics. Hence, similar to other countries such as the USA, a Twitter analysis could be used for retracing political events and social movements.

The qualitative analysis of political topics suggests that Nigerians tweet mostly about recent or upcoming political events. Since we detected at least one topic related to offline protests for each considered month, we expect other months to contain topics related to protests as well. The qualitative analysis of the most important topic words suggests that the protest-related topics focus on the protest cause rather than on the protest itself. Hence, identifying topics related to Nigerian protests is not a straightforward task but instead requires a comparison between the topics' content and the protest cause.

5.2 Automatized Tweet-Protest Mapping

The mapping quality assessment indicates that the most simple mapping approach performs best, while the more complicated embedding approaches fall behind. An explanation for the superior performance of the hashtag string comparison could be that a couple of buzzwords extracted from hashtags by identifying relevant subwords convey only the core information of the movement. Thus, if the movement is related to a protest, this core information is also used in most cases for describing the related protest. Hence, movement hashtags contain just enough information to identify the core message while avoiding being too general.

All other mappings use more information on the movements' content. However, this added information seems to generalize the information, which leads to multiple overlapping content words and embeddings. For instance, three roughly simultaneous movements could talk about violence against women. One keeps the discussion on a general level while one focuses on the rape of women and the other on the killings of women. By taking more content information of these movements into account, the difference in the form of violence could fade into the background making a precise mapping more complex.

This difficulty caused by too general information applies not only to the movement information but also to the protest information, which can be observed when comparing the two Glove Embedding approaches. While the mean approach performs worst, the word-based similarity approach performs second best. The mean embeddings of protest notes yield mappings to general movements instead of to smaller specific movements. Hence, protests that are not similar to a specialized movement are mapped to general movements, leading to a low recall score. In contrast, the word-based similarity score can identify buzzwords that stand out for only one or a few movements. Hence, similar to the hashtag string comparison, this approach can identify specific enough information for accurate mapping.

In general, all approaches are limited by the small amount of information contained in the protest notes of the ACLED data set.

Achieving high-performance metrics is especially hard for this setting since we want to not only discover whether a protest was discussed online but also correctly map the movement related to the protest. Hence, in addition to protests mapped to a movement manually but not by our mapping approach, the false-positive class contains protests mapped to incorrect movements. For instance, the two movements '#revolutionnow' and '#freesowore' are strongly related. '#revolutionnow' is the superior movement fighting against bad governance, while '#freesowore' is about freeing the arrested leader Omoyele Sowore of the '#revolutionnow' movement. If our approach maps a protest to the '#revolutionnow' movement instead of the strongly related '#freesowore' movement, it is still registered as false-positive. Hence, achieving a high true-positive value is more challenging than a high true-negative value. Along with the positive class being the minority, this explains the higher recall compared to the precision scores for most approaches.

Further, we restricted the manual labeling to hashtags used in a protest context only. However, we observed that protests have an additional impact on more general hashtags in multiple cases. For instance, protests in the context of the Nigerian election in 2019 affected the general '#nigeriadecides2019' hashtag. If no more specific movement exists, most of our mapping approaches can capture this relation. Consequently, these additional mappings lead to multiple false-negative mappings.

5.3 Inter-Protest-Analysis

In this section, we discuss and interpret the results of the inter-protest analysis. Initially, in Section 5.3.1, we interpret the identified categories of movement-protest mappings by setting the results in a real-world context as well as in a social media research context. In Section 5.3.2, we discuss the behavior of the individual features of short-term movements. Subsequently, we interpret the comparison results of individual short-term movements and short-term sub-movements of long-term movements in Section 5.3.3. Lastly, we discuss the results of analyzing the mapped protests in Section 5.3.4

5.3.1 Categories of Movement-Protest Mappings

Results of the duration analysis show a clear division at roughly three months between short-term and long-term movements.

The examples for short-term Twitter volume trends shown in Figure 4.3 follow roughly the typical behavior of exogenous trends, i.e., trends triggered by external events [97]: a sudden burst of activity is followed by a power-law relaxation. Further, the short-term movement duration between a few days and a couple of months also complies with the results of exogenous trend analysis on Twitter [98]. In contrast, the Twitter volume trend of the long-term sample movement in Figure 4.3d indicates that long-term movements are not constantly active but rather emerge multiple times with extended inactive periods in between.

Depending on the positional relation between Twitter volume peak day and protest day, the nature of the external events triggering an online trend differs. Movements that peak before the protest day (Category ST_I) are triggered by an event a few days before the emergence. For instance, the '#saynotosocialmediabill' movement shown in Figure 4.3a starts emerging the day after the second senate reading of the controversial social media bill. Hence, these movements aim to show their disapproval of a past event. Movements that peak on the protest day (Category ST_II) are triggered by a planned event happening on the protest day. These movements either aim to celebrate annual holidays such as International Women's Day or to oppose or support a planned event. For instance, the '#ngobill' movement shown in Figure 4.3b took place in front of the national assembly during a public hearing in order to prevent the passage of the bill. In contrast, movements emerging shortly after a protest (Category ST_III) are typically triggered by this protest. Qualitative analysis of the protest notes indicates that typically during these protests, extensive violence was used - often initiated by police forces. For instance, the '#justiceforfuoye' movement shown in Figure 4.3c was triggered by the death of a student killed by policemen during a protest at Fuoye University. Hence, these movements arise from the public's disapproval of the used violence during the protest.

5.3.2 Characteristics of Short-Term Movements

In the following, we discuss the observed behavior of computed features presented in Section 4.3 and the dependencies between them.

Twitter Volume. The results of the Twitter volume trend analysis indicate that online movements follow roughly the behavior of typical externally triggered trends [98, 97]. However, the movement's decline is slightly interrupted on the day of the protest, just to fade away shortly afterward. Hence, the protest does not act as a catalyst for another movement's rise but only delays its fading.

The observed variables related to the Twitter volume show dependencies between each other. The higher the Twitter volume peak, the more people are involved in the movement. Collective actions involving more people require more organization. Hence, the time period between the movement's emergence and the protest is longer. If the protest happens shortly after the Twitter volume peak, the movement is still at the beginning of its decline. Hence, the effect on the Twitter volume trend of the beginning decline and the protest overlap. Consequently, more people are still active in the movement's discussion on the day of the protest, leading to a larger ratio of the Twitter volume on the protest day than on the peak day. In conclusion, the dependency between Δ and $n_{protest}/n_{peak}$ evolves from the interplay of the short attention span online and the attention for real-world collective actions on social media.

User Engagement. In general, the mapped movements exhibit the highest user engagement during the movement's emergence until reaching the Twitter volume peak. Similar to the Twitter volume, the trend on the protest day for both user activity and the fraction of new movement participants indicates that a protest keeps the movement from diminishing rather than triggering another rise. In conclusion, the peak in Twitter volume on the protest day arises from the increased user activity of movement participants on that day and not from a more extensive spread of the movement.

Regarding the interaction via annotations, we discuss the results of the social network analysis in the following. The hierarchical structure shows that the movement information is distributed in levels. If applicable, the initial layer consists of a major organization that initiated the online movement and the protest, for instance, the Twitter account of 'change.org'. The next level contains official Twitter accounts, including politicians, NGOs, state organizations, and activists. This layer is followed by layers of Twitter accounts that are highly active in the respective movement. They convey the message of the organizers, spread the information by annotating other accounts, and draw the attention of officials by annotating them. The last layers consist of Twitter accounts that rely on the information of the layers with the most active accounts. Hence, the initial layers of hierarchy typically include the Twitter accounts of entities whose attention should be drawn to the matter of the movement. In contrast, the lower layers contain Twitter accounts of members of the general population who are trying to be mobilized by others. The results of the centrality measure reinforce this insight. The indegree centrality analysis indicates that the online discussions are quickly escalated to a couple of the most powerful Nigerian politicians, including the current and former Nigerian presidents and vice-presidents. Further, the changes in the distribution of indegree centrality on the protest day can be explained by the attention to collective actions of the general public. Social media accounts of news channels report about protests taking place, which is why they are mentioned more frequently by other accounts on that day. In addition, some Nigerian celebrities habitually publicly take a stand on the protest matter, which also leads to increased mentioning behavior. Lastly, the diverse nature of the outdegree centrality indicates that the Twitter accounts actively spreading the information to selected others by annotations differ strongly across protests.

Mood of Twitter Discussion. In contrast to a study on social media sentiment for a protest in India, [19], the observed Nigerian movements show a neutral-negative sentiment. There are multiple potential explanations for this substantial deviation that could interplay: First, the protests and the online movements happened in different countries with different cultural characteristics that can influence the nature of discussions. Second, the sentiment classification methods have improved significantly since the research was published. Third, Twitter became known for its negative discussion climate over the years. Hence, discussions nowadays could be more negative than a few years ago.

Development over the Years. During the first four considered years (2016-2019), the nature of the online movements changed only slightly. However, in the year 2020, the change is significant. The number of "freshly" created users suggests that spam and fake accounts reached Nigeria in 2020, which drastically distorts the online movements' behavior. Further, in 2020 multiple international movements and global topics, including Covid-19, spread to Nigeria. Related protests took place in various Nigerian cities. Hence, the mapped movements became less national and more international, which is another explanation for this drastic change in 2020.

5.3.3 Short-Term Sub-Movements in Long-Term Movements

The comparison between short-term movements and short-term sub-movements, which are part of a superior long-term movement, indicates a couple of differences.

Foremost, long-term movements have already established an active and often passionate user base. Thus, in comparison to short-term movements, the user activity for short-term sub-movements is higher, while the fraction of new movement participants is lower. However, this behavior depends on the positioning of the sub-movement within the long-term movement. If the long-term movement only just emerged for the first time or was not active for an extended period, the behavior is more similar to an individual short-term movement. Due to this behavior variation, the standard deviation measured for variables of short-term sub-movements is significantly higher.

Lastly, long-term movements last long enough in order to gain international attention. Hence, the ratio of Nigerian Twitter accounts is lower for short-term sub-movements than for the individual, mostly regional, short-term movements.

5.3.4 Analysis of Mapped Protests

In general, the results of analyzing the properties of protests mapped by our approach indicate that Twitter usage in the context of protests increased significantly in Nigeria over five years. This insight is consistent with the ongoing growth of Social Media in Nigeria.

The distribution of protest types among the mapped protests compared to all registered protests suggests that involved violence leads to more reliable online discussions. Violent acts, especially if executed by police forces, trigger a groundswell of disapproval among the Nigerian population. Hence, even small protests are discussed intensely online if they involve the usage of the police force. This emergence of social movements as disapproval of violence after such protests could explain their reliable mapping to online social movements.

The geographic distribution analysis results indicate that Twitter as a communication tool for social movements is mainly used by the southern population of Nigeria. This observation agrees with Nigeria's economic and social imbalances between the more developed south and the rural north.

5.4 Limitations and Future Work

The quantitative results and their interpretations, presented in the previous sections, have to be seen in light of some limitations regarding the employed data sources and the applied methods.

Data sets. In general, we restrict all data sets to the English language. Since English is the official language in Nigeria, this is a valid choice. However, in more rural areas of Nigeria, people use stronger dialects and even different languages. Hence, tweets by these people in their mother tongue are disregarded. In addition, the completeness of the data sets is limited by the employed scraping tool TWINT. Since TWINT cannot scrape retweets, we have to disregard this fraction of tweets and interactions.

As a basis for this research, we employed a data set of tweets geotagged with locations in Nigeria. During the inter-protest analysis, we observed that the ratio of geotagged tweets within the movements is vanishingly small («5%). In conclusion, Nigerians tend not to use the geotag option in their tweets which sheds light on a limitation of the geo-referenced data set: Due to the rare use of geotags, relevant hashtags could be missing in the geo-referenced data set.

Further, the data set employed for identifying online movements contains only tweets that use a specific hashtag. Since not all tweets related to a movement include hashtags, our approach cannot extract them. However, without considering further context information, such as who is following who, it is not possible to capture these tweets without hashtags reliably.

Lastly, the protest data set brings some limitations due to their methods of detecting protests via conventional data sources and the information they register. For instance, the protest note is manually entered by the registrant and differs slightly across protests. Reasons for these deviations could be simply daily variations of the registrant and the use of different registrants.

Methods. Regarding the methodology, each research part brings its limitations. In general, all applied NLP models are pre-trained on English data. Although Nigeria's official language is English, they developed a dialect that differs in some aspects from standard English. This deviation can limit the NLP models' performances.

We employ a pipeline of different methods for the automatized mapping of tweets to protest. Hence, the limitations of the initial methods impact the performance of the last methods. For instance, if the topic similarity approach does not identify relevant hashtags, it is not included in the set of movements and thus, can not be mapped to any protest. Since we only assess the quality of the final mappings, we can not conclude how much impact the limitations of each pipeline step have on the final mapping quality. Performing an ablation study could shed light on the nature of these limitations.

The analysis identifies typical behavior across movement-protest mappings using the available information. For daily aggregations, we omit all days with too few data points. Still, the number of data points per day differs strongly. Thus, the reliability of the aggregated values deviates. Further, we analyze dependencies between various variables. However, we can only examine correlation and not causation with our analysis results. In addition, all insights focus on a particular type of movement-protest mappings. Although this type is most frequent among all mappings, we disregard a fraction of movement-protest mappings in our analysis. Hence, the analysis results can not be transferred to all movements and protests in Nigeria. Further, in this research, we focus on Nigeria only. Thus, it is unclear how the insights are transferable to other countries. Lastly, we decided to omit other context information such as political stability in Nigeria, climate conditions, and economic conditions. These context settings can affect the behavior of people offline and online. However, since we aggregate across multiple years in our analysis, we expect a change in these conditions to have only a small impact on the aggregated features.

Future Work. Based on our results and the limitations, we identify multiple directions for future research in this area.

First, comparing online movements related to protests across countries could lead to valuable insights into the similarities and differences of protest-culture across countries. The approaches introduced in this thesis are all easily adaptable to other countries. Further, all data sets are either available or contractible for other countries. Hence, the major challenge in transferring our methods to other countries could be the language restriction of some NLP models.

Further, the analysis could compare online movements that lead to protests with online movements that do not. Since we extracted a large set of bursting movements, this could be done by comparing characteristics of mapped and unmapped movements. This comparison could clarify why some online social movements translate into offline actions and some not.

Using this knowledge, the computed features can be used for predicting protests. In terms of protest prediction we initially tested some complex prediction models using Longformers [99] on tweets and Temporal Graph Neural Networks [100] on daily interaction networks. Both methods failed to learn the relationship between tweets and protests in a straightforward manner. In addition, we observed that for the movement-protest mappings more conventional and less complex methods performed better on this kind of noisy user-generated data. Hence, using the gained knowledge of this analysis and applying simple statistical models such as SVMs or simple neural networks on the relevant characteristics could be a promising approach for protest prediction.

In general, predicting the date and the nature of protests in advance can help to provide appropriate safety measures. However, such prediction tasks also need to be carefully evaluated to prevent discriminative biases and misuse by undemocratic regimes.

Chapter 6

Conclusion

This thesis aimed to investigate the nature of relations between online social movements and protests in Nigeria. We extracted and explored relevant data, introduced an approach for mapping online social movements to protests and performed a quantitative comparative analysis across hundreds of online social movements and protest pairs.

Constructing a Twitter data set consisting of Nigerian tweets only enabled us to explore the Twitter landscape of one country. Using this data set, we modelled topics via a BERT-based clustering approach. We identified politics as one prominent subject among these topics and came across various protest related political topics. Hence, political discourse accounts for a non-negligible part of Nigeria's Twitter landscape. The political discourse reaches from discussion of simultaneously happening offline political events to discussing future offline events and planning collective actions.

We took the presence of protest related topics as opportunity to introduce an automated approach for detecting relevant online social movements and mapping them to related protests. Building upon hashtag activism, the approach identifies potentially relevant hashtags, extracts data from Twitter, filters bursting hashtags and clusters the collected data into online social movements. Further, we proposed and assessed various techniques for mapping online social movements to related protests. Here, we observed that comparing hashtag words and protest information outperforms more complex mapping techniques.

We performed a quantitative comparative analysis based on the mappings of online social movements and protests. We computed features for capturing Twitter volume, user engagement and mood of the discussions by employing various data analysis approaches including social network analysis and sentiment classification. By comparing basic characteristics of these features, we determined multiple types of relations between online social movements and protests. Based on the subsequent comparative analysis of one relation type, we identified differing characteristics and formulated and interpreted common trends of protest-related online social movements in Nigeria.

In conclusion, this thesis uses state-of-the art Big Data analysis techniques to identify and analyse relevant characteristics that determine the relation between online social movements and protests in Nigeria. We introduced an approach for performing quantitative comparative analyses for protest-related social movements using social media data. Due to its modular nature, this approach could be easily transferred to other types of offline social events. Consequently, this work provides a basis for future research on event analysis, the translation of online action to offline action, and protest prediction.

Bibliography

[1] P. N. Howard, A. Duffy, D. Freelon, M. M. Hussain, W. Mari, and M. Maziad, "Opening closed regimes: what was the role of social media during the arab spring?" *Available at SSRN 2595096*, 2011.

[2] C. Fuchs, "Social media, riots, and revolutions," *Capital & Class*, vol. 36, no. 3, pp. 383–391, 2012.

[3] K. M. DeLuca, S. Lawson, and Y. Sun, "Occupy wall street on the public screens of social media: The many framings of the birth of a protest movement," *Communication, Culture & Critique*, vol. 5, no. 4, pp. 483–509, 2012.

[4] T. B. I. for Global Change. (2021) Social media futures: How to change the african narrative. [Online]. Available: https://institute.global/policy/social-media-futures-how-change-african-narrative

[5] Twitter. (2021) Establishing twitter's presence in africa. [Online]. Available: https://blog.twitter.com/en_us/topics/company/2021/establishing-twitter-s-presence-in-africa

[6] CNN. (2021) Nigeria bans twitter after company deletes president buhari's tweet. [Online]. Available: https://edition.cnn.com/2021/06/04/africa/nigeria-suspends-twitter-operations-intl/index.html

[7] T. World. (2021) Why has nigeria pulled the plug on twitter? [Online]. Available: https://www.trtworld.com/magazine/why-has-nigeria-pulled-the-plug-on-twitter-47352

[8] A. International. (2021) endsars movement: From twitter to nigerian streets. [Online]. Available: https://www.amnesty.org/en/latest/campaigns/2021/02/nigeria-end-impunity-for-police-violence-by-sars-endsars

[9] BBC. (2021) Nigeria's #endsars protests: What happened next. [Online]. Available: https://www.bbc.com/news/world-africa-58817690

[10] Aljazeera. (2019) Nigerians raise alarm over controversial social media bill. [Online]. Available: https://www.aljazeera.com/news/2019/12/18/nigerians-raise-alarm-over-controversial-social-media-bill

[11] C. TV. (2021) Twitter's mission in nigeria is suspicious, says lai mohammed. [Online]. Available: https://www.channelstv.com/2021/06/02/twitters-mission-in-nigeria-is-suspicious-says-lai-mohammed/

[12] BBC. (2021) Twitter ban in nigeria: President buhari lift ban on twitter afta seven months. [Online]. Available: https://www.bbc.com/pidgin/59969898

[13] J. Ince, F. Rojas, and C. A. Davis, "The social media response to black lives matter: How twitter users interact with black lives matter through hashtag use," *Ethnic and racial studies*, vol. 40, no. 11, pp. 1814–1830, 2017.

[14] A. Karduni and E. Sauda, "Anatomy of a protest: spatial information, social media, and urban space," *Social Media+ Society*, vol. 6, no. 1, p. 2056305119897320, 2020.

[15] M. T. Bastos, R. Recuero, and G. Zago, "Taking tweets to the streets: A spatial analysis of the vinegar protests in brazil," *First Monday*, vol. 19, no. 3, 2014.

[16] J. M. Larson, J. Nagler, J. Ronen, and J. A. Tucker, "Social networks and protest participation: Evidence from 130 million twitter users," *American Journal of Political Science*, vol. 63, no. 3, pp. 690–705, 2019.

[17] V. Spaiser, T. Chadefaux, K. Donnay, F. Russmann, and D. Helbing, "Communication power struggles on social media: A case study of the 2011–12 russian protests," *Journal of Information Technology & Politics*, vol. 14, no. 2, pp. 132–153, 2017.

[18] A. Sinpeng, "Hashtag activism: social media and the# freeyouth protests in thailand," *Critical Asian Studies*, vol. 53, no. 2, pp. 192–205, 2021.

[19] A. S. Neogi, K. A. Garg, R. K. Mishra, and Y. K. Dwivedi, "Sentiment analysis and classification of indian farmers' protest using twitter data," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100019, 2021.

[20] A. Zervopoulos, A. G. Alvanou, K. Bezas, A. Papamichail, M. Maragoudakis, and K. Kermanidis, "Hong kong protests: Using natural language processing for fake news detection on twitter," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2020, pp. 408–419.

[21] B. N. Hafez, "New social movements and the egyptian spring: a comparative analysis between the april 6 movement and the revolutionary socialists," *Perspectives on Global Development and Technology*, vol. 12, no. 1-2, pp. 98–113, 2013.

[22] D. M. Clayton, "Black lives matter and the civil rights movement: A comparative analysis of two social movements in the united states," *Journal of Black Studies*, vol. 49, no. 5, pp. 448–480, 2018.

[23] C. C. Mann, "Choosing an indigenous official language for nigeria." 1990.

[24] I. M. Fund. (2022) World economic outlook database. [Online]. Available: https://www.imf.org/en/Publications/WEO/weo-database

[25] F. Baumann, "The next frontier—human development and the anthropocene: Undp human development report 2020," *Environment: Science and Policy for Sustainable Development*, vol. 63, no. 3, pp. 34–40, 2021.

[26] T. International. (2021) Corruption perceptions index 2020. [Online]. Available: https://www.transparency.org/en/cpi/2021

[27] I. P. S. Association. (2016) World internal security & police index (wispi). [Online]. Available: http://www.ipsa-police.org/images/uploaded/Pdf%20file/WISPI%20Report.pdf

[28] H. R. M. Initiative. (2018) Nigeria - hrmi rights tracker. [Online]. Available: https://rightstracker.org/en/country/NGA

[29] Aljazeera. (2019) Death, anguish and flickers of hope: 10 years of boko haram. [Online]. Available: https://www.aljazeera.com/news/2019/7/30/death-anguish-and-flickers-of-hope-10-years-of-boko-haram

[30] R. M. Olapeju and A. O. Peter, "The impact of banditry on nigeria's security in the fourth republic: An evaluation of nigeria's northwest," *Zamfara Journal of Politics and Development*, vol. 2, no. 1, pp. 26–26, 2021.

[31] A. International. (2021) Nigeria: At least 115 people killed by security forces in four months in country's southeast. [Online]. Available: https://www.amnesty.org/en/latest/news/2021/08/nigeria-at-least-115-people-killed-by-security-forces-in-four-months-in-countrys-southeast/

[32] U. Egbuta, "Understanding the herder-farmer conflict in nigeria," *Conflict Trends*, vol. 2018, no. 3, pp. 40–48, 2018.

[33] A. International, *'You Have Signed Your Death Warrant': Torture and Other Ill Treatment by Nigeria's Special Anti-robbery Squad (SARS)*. Amnesty International, 2016.

[34] A. International. (2020) Nigeria: Horrific reign of impunity by sars makes mockery of anti-torture law. [Online]. Available: https://www.amnesty.org/en/latest/news/2020/06/nigeria-horrific-reign-of-impunity-by-sars-makes-mockery-of-anti-torture-law

[35] T. Guardian. (2014) Bring back our girls: global protests over abduction of nigerian schoolgirls. [Online]. Available: https://www.theguardian.com/world/2014/may/07/bring-back-our-girls-global-protests-abduction-nigerian-schoolgirls

[36] P. Gazette. (2022) Days after bandits killed over 200 in zamfara, northern group says "north not bleeding". [Online]. Available: https://gazettengr.com/days-after-bandits-killed-over-200-in-zamfara-northern-group-says-north-not-bleeding/

[37] F. Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," Cornell Aeronautical Lab Inc Buffalo NY, Tech. Rep., 1961.

[38] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[39] Y. Goldberg, "Neural network methods for natural language processing," *Synthesis lectures on human language technologies*, vol. 10, no. 1, pp. 1–309, 2017.

[40] J. L. Elman, "Finding structure in time," *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.

[41] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[42] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[43] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[44] T. Rocktäschel, E. Grefenstette, K. M. Hermann, T. Kočiskỳ, and P. Blunsom, "Reasoning about entailment with neural attention," *arXiv preprint arXiv:1509.06664*, 2015.

[45] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[46] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[47] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.

[48] J. Cheng, L. Dong, and M. Lapata, “Long short-term memory-networks for machine reading,” *arXiv preprint arXiv:1601.06733*, 2016.

[49] R. Baeza-Yates, B. Ribeiro-Neto *et al.*, *Modern information retrieval.* ACM press New York, 1999, vol. 463.

[50] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to information retrieval.* Cambridge University Press Cambridge, 2008, vol. 39.

[51] J. Pennington, R. Socher, and C. D. Manning, “Glove: Global vectors for word representation,” in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[52] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *arXiv preprint arXiv:1301.3781*, 2013.

[53] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching word vectors with subword information,” *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017.

[54] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, “Deep contextualized word representations,” 2018. [Online]. Available: http://arxiv.org/abs/1802.05365

[55] A. Akbik, D. Blythe, and R. Vollgraf, “Contextual string embeddings for sequence labeling,” in *Proceedings of the 27th international conference on computational linguistics*, 2018, pp. 1638–1649.

[56] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey *et al.*, “Google’s neural machine translation system: Bridging the gap between human and machine translation,” *arXiv preprint arXiv:1609.08144*, 2016.

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

[58] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, “Tweeteval: Unified benchmark and comparative evaluation for tweet classification,” *arXiv preprint arXiv:2010.12421*, 2020.

[59] N. Reimers and I. Gurevych, “Sentence-bert: Sentence embeddings using siamese bert-networks,” *arXiv preprint arXiv:1908.10084*, 2019.

[60] J. L. Boyd-Graber, Y. Hu, D. Mimno *et al.*, *Applications of topic models.* now Publishers Incorporated, 2017, vol. 11.

[61] H. Zhao, D. Phung, V. Huynh, Y. Jin, L. Du, and W. Buntine, "Topic modelling meets deep neural networks: A survey," *arXiv preprint arXiv:2103.00498*, 2021.

[62] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[63] C. Févotte and J. Idier, "Algorithms for nonnegative matrix factorization with the β-divergence," *Neural computation*, vol. 23, no. 9, pp. 2421–2456, 2011.

[64] P. Xie and E. P. Xing, "Integrating document clustering and topic modeling," *arXiv preprint arXiv:1309.6874*, 2013.

[65] S. Sia, A. Dalmia, and S. J. Mielke, "Tired of topic models? clusters of pretrained word embeddings make for fast and good topics too!" *arXiv preprint arXiv:2004.14914*, 2020.

[66] F. Viegas, S. Canuto, C. Gomes, W. Luiz, T. Rosa, S. Ribas, L. Rocha, and M. A. Gonçalves, "Cluwords: exploiting semantic word clustering representation for enhanced topic modeling," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, 2019, pp. 753–761.

[67] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering." *J. Open Source Softw.*, vol. 2, no. 11, p. 205, 2017.

[68] D. Angelov, "Top2vec: Distributed representations of topics," *arXiv preprint arXiv:2008.09470*, 2020.

[69] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based tf-idf procedure," *arXiv preprint arXiv:2203.05794*, 2022.

[70] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[71] W. Dou, X. Wang, W. Ribarsky, and M. Zhou, "Event detection in social media data," in *IEEE VisWeek workshop on interactive visual text analytics-task driven analytics of social media content*, 2012, pp. 971–980.

[72] S. Petrović, M. Osborne, and V. Lavrenko, "Streaming first story detection with application to twitter," in *Human language technologies: The 2010 annual conference of the north american chapter of the association for computational linguistics*, 2010, pp. 181–189.

[73] R. Li, K. H. Lei, R. Khadiwala, and K. C.-C. Chang, "Tedas: A twitter-based event detection and analysis system," in *2012 IEEE 28Th international conference on data engineering.* IEEE, 2012, pp. 1273–1276.

[74] F. Zarrinkalam and E. Bagheri, "Event identification in social networks," *Encyclopedia with Semantic Computing and Robotic Intelligence*, vol. 1, no. 01, p. 1630002, 2017.

[75] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 851–860.

[76] A.-M. Popescu and M. Pennacchiotti, "Detecting controversial events from twitter," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, 2010, pp. 1873–1876.

[77] E. Benson, A. Haghighi, and R. Barzilay, "Event discovery in social media feeds," in *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*, 2011, pp. 389–398.

[78] R. Lee and K. Sumiya, "Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection," in *Proceedings of the 2nd ACM SIGSPATIAL international workshop on location based social networks*, 2010, pp. 1–10.

[79] K. C. Roy, M. Cebrian, and S. Hasan, "Quantifying human mobility resilience to extreme events using geo-located social media data," *EPJ Data Science*, vol. 8, no. 1, p. 18, 2019.

[80] J. Ansah, L. Liu, W. Kang, J. Liu, and J. Li, "Leveraging burst in twitter network communities for event detection," *World Wide Web*, vol. 23, no. 5, pp. 2851–2876, 2020.

[81] P. Tijare and J. R. Prathuri, "A survey on event detection and prediction online and offline models using social media platforms," *Materials Today: Proceedings*, 2021.

[82] M. Tremayne, "Anatomy of protest in the digital era: A network analysis of twitter and occupy wall street," *Social movement studies*, vol. 13, no. 1, pp. 110–126, 2014.

[83] P. S. Lee, C. Y. So, and L. Leung, "Social media and umbrella movement: Insurgent public sphere in formation," *Chinese Journal of Communication*, vol. 8, no. 4, pp. 356–375, 2015.

[84] A. Scherman and S. Rivera, "Social media use and pathways to protest participation: Evidence from the 2019 chilean social outburst," *Social Media+ Society*, vol. 7, no. 4, p. 20563051211059704, 2021.

[85] S. Valenzuela, T. Correa, and H. Gil de Zuniga, "Ties, likes, and tweets: Using strong and weak ties to explain differences in protest participation across facebook and twitter use," *Political communication*, vol. 35, no. 1, pp. 117–134, 2018.

[86] C. Raleigh, A. Linke, H. Hegre, and J. Karlsen, "Introducing acled: an armed conflict location and event dataset: special data feature," *Journal of peace research*, vol. 47, no. 5, pp. 651–660, 2010.

[87] T. A. C. Location and E. D. Project. (2017) Armed conflict location event data project (acled) codebook. [Online]. Available: https://www.acleddata.com/wp-content/uploads/dlm_uploads/2017/10/ACLED_Codebook_2019FINAL_pbl.pdf

[88] T. A. C. Location and E. D. Project. (2019) Acled coverage to date. [Online]. Available: https://acleddata.com/acleddatanew/wp-content/uploads/dlm_uploads/2019/01/ACLED_Country-and-Time-Period-coverage_updatedFeb2022.pdf

[89] (2018) Twint - twitter intelligence tool. [Online]. Available: https://github.com/twintproject/twint

[90] S. Schlosser, D. Toninelli, and M. Cameletti, "Comparing methods to collect and geolocate tweets in great britain," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 1, p. 44, 2021.

[91] Nigeria cities database. [Online]. Available: https://simplemaps.com/data/ng-cities

[92] relatedwords. [Online]. Available: https://relatedwords.org

[93] 100m characters of wikipedia. [Online]. Available: http://mattmahoney.net/dc/text8.zip

[94] J. Kleinberg, "Bursty and hierarchical structure in streams," *Data mining and knowledge discovery*, vol. 7, no. 4, pp. 373–397, 2003.

[95] J. Luo and C. L. Magee, "Detecting evolving patterns of self-organizing networks by flow hierarchy measurement," *Complexity*, vol. 16, no. 6, pp. 53–61, 2011.

[96] U. Brandes and D. Wagner, "Analysis and visualization of social networks," in *Graph drawing software*. Springer, 2004, pp. 321–340.

[97] R. Crane and D. Sornette, "Robust dynamic classes revealed by measuring the response function of a social system," *Proceedings of the National Academy of Sciences*, vol. 105, no. 41, pp. 15 649–15 653, 2008.

[98] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 591–600.

[99] I. Beltagy, M. E. Peters, and A. Cohan, "Longformer: The long-document transformer," *arXiv preprint arXiv:2004.05150*, 2020.

[100] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," *arXiv preprint arXiv:2006.10637*, 2020.