POLITECNICO
MILANO 1863

SCHOOL OF CIVIL
ENVIRONMENTAL
AND LAND MANAGEMENT
ENGINEERING

Spatio-temporal Dynamics and Network Diffusion of Weibo Sentiment Toward U.S. Tariff Policies: A Geographic Perspective

Master of Science Degree in
Geoinformatics Engineering

Author: **Jianwei Deng**

Student ID: 245188
Advisor: Prof. Francesco Pierri
Co-advisors: Geng Liu
Academic Year: 2025-26

Abstract

Global policy shocks trigger complex waves of public sentiment that vary across regions and social groups. However, the mechanism by which regional economic sensitivities and social influence structures jointly shape these responses remains poorly understood. Focusing on United States tariff policies in early 2025, this study examines how event driven shocks across various sectors trigger sentiment shifts and information diffusion within the Weibo ecosystem.

We constructed a large scale dataset comprising 1.3 million posts, including 230,000 geolocated entries. To enable high precision analysis, we implemented a weakly supervised learning framework by fine tuning a specialized RoBERTa model using pseudo labels generated by DeepSeek-v3. Our analytical framework integrates temporal dynamics, Geographically Weighted Regression (GWR) to capture spatial heterogeneity, and information cascade modeling to evaluate network diffusion.

The results reveal a structured societal response. Temporally, sentiment follows a multi peak evolution across three strategic policy stages, where event cycles are dynamically defined using moving average smoothing and a 20% peak intensity threshold. Spatially, GWR identifies a critical polarization phase where local economic factors explain up to 57.2% of sentiment variance. Topologically, we observe a structural divergence in propagation: institutional accounts (Blue V) drive explosive breadth, whereas elite influencers and regular users act as the primary drivers of multi level cascade depth. These findings quantify how digital public opinion is co-constructed by geographic economic realities and social network hierarchies.

Keywords: Computational Social Science, Information Diffusion, Spatiotemporal Dynamics, Tariff, US China Trade Conflict, Large Language Models, Weibo.

Abstract in lingua italiana

Gli shock legati alle politiche globali scatenano complesse ondate di opinione pubblica, ma l'interazione tra sensibilità economiche regionali e strutture di influenza sociale è ancora poco compresa. Focalizzandosi sulle politiche tariffarie degli Stati Uniti all'inizio del 2025, questo studio esamina come gli shock guidati dagli eventi in vari settori inneschino cambiamenti nel sentimento e la diffusione delle informazioni all'interno dell'ecosistema di Weibo.

Abbiamo costruito un set di dati su larga scala di 1,3 milioni di post, inclusi 230.000 record geolocalizzati, e implementato un framework di apprendimento debolmente supervisionato perfezionando un modello RoBERTa specializzato tramite etichette generate da DeepSeek-v3. L'analisi integra dinamiche temporali, regressione geograficamente pesata (GWR) per catturare l'eterogeneità spaziale e modellazione delle cascate di informazioni per valutare la diffusione di rete.

I risultati rivelano una risposta sociale strutturata. Temporalmente, il sentimento segue un'evoluzione a picchi multipli attraverso tre fasi politiche strategiche, con cicli di eventi definiti dinamicamente tramite medie mobili e una soglia di intensità del 20% rispetto al picco. Spazialmente, la GWR identifica una fase di polarizzazione critica in cui i fattori economici locali spiegano fino al 57,2% della varianza del sentimento. Topologicamente, emerge una divergenza strutturale nella propagazione: gli account istituzionali (Blue V) guidano l'ampiezza esplosiva, mentre gli influencer e gli utenti comuni sono i motori primari della profondità delle cascate multilivello.

Parole chiave: Scienza sociale computazionale, Diffusione delle informazioni, Dinamiche spaziotemporali, Tariffe, Conflitto commerciale USA Cina, Modelli linguistici di grandi dimensioni, Weibo.

Contents

Introduction

Monitoring public sentiment toward macroeconomic shifts has become increasingly critical as global trade volatility accelerates [1]. Social media platforms provide high-frequency behavioral data that complements traditional economic indicators, which often suffer from reporting lags [2]. Among these digital sources, Sina Weibo serves as a primary social sensorium in China, enabling the observation of collective emotional responses to international policy shocks such as the 2025 U.S. tariff impositions [3, 4].

However, the effective analysis of these data is limited by the linguistic complexity of policy-related discourse and the geographic non-stationarity of economic impacts [5]. Most existing sentiment models fail to account for the intentional and sequential nature of trade shocks, leading to a lack of granularity in understanding how regional economic dependencies dictate public reaction [6]. Furthermore, the mechanisms of information diffusion in moderated environments remain under-explored, particularly regarding how legitimate policy sentiment propagates across different user strata compared to misinformation [7, 8]. The 2025 tariff surge—an aggressive exogenous shock affecting China's export-oriented provinces—offers a compelling case to quantify the interplay between regional economic sensitivity and digital influence structures [9].

This research develops a multi-dimensional framework to analyze 1.3 million Weibo posts and 230,000 geolocated entries. The approach implements a weakly supervised learning pipeline that distills knowledge from the DeepSeek-v3 Large Language Model into a specialized RoBERTa classifier for high-precision sentiment detection [10]. We integrate Geographically Weighted Regression (GWR) to link spatial sentiment variance with provincial trade-to-GDP ratios [11], and utilize information cascade modeling to evaluate the structural virality of propagation across institutional and individual accounts [12, 13].

The thesis quantifies temporal sentiment evolution across dynamic event cycles, details the spatial heterogeneity of public concern over industrial hubs, and evaluates the diffusion efficiency of five distinct user categories. The resulting framework aims to facilitate a more reliable and explanatory use of social media data for monitoring the societal impact of global macroeconomic policies.

1 | Background And Motivation

1.1. Research Background

The trade conflict between the United States and China has evolved from targeted measures into a systemic economic decoupling. By early 2025, this confrontation intensified under the second Trump administration with the imposition of a universal 10% tariff on Chinese goods (effective February 4, 2025) and reciprocal adjustments under Executive Order 14257 (April 2, 2025). These policy interventions, directly impacting over US$550 billion of annual exports, serve as exogenous shocks that trigger profound waves of public sentiment and information diffusion within Chinese society.

In this context, Sina Weibo serves as a critical sensor for collective societal response [14]. With over 600 million monthly active users, it constitutes a unique information ecosystem shaped by specific algorithmic mechanisms and user strata [17]. This platform provides an ideal environment to study how political expression and collective emotions propagate in response to external economic pressures within a moderated public sphere [4].

1.2. Research Questions

The central research question addressed in this thesis is:

> *Quantify the response patterns of Weibo users to tariff policies,*
> *Uncover the inner mechanisms of these responses in terms of time, space and interpersonal networks.*

This is further supported by the following sub-questions from temporal, spatial, and network topology dimensions:

- **RQ1:** How do tariff-related shocks shape the temporal evolution of public sentiment?

- **RQ2:** How do sentiment and topics differ across regions, and what are the spatial patterns of these reactions?
- **RQ3:** How does user membership influence the breadth and depth of sentiment diffusion in retweet networks?

1.3. Objectives

By addressing these questions, this thesis aims to achieve the following core objectives:

1. **Quantify Temporal Sentiment Shifts:** To measure how emotional intensity and polarity fluctuate in direct response to specific policy announcement windows.
2. **Analyze Spatial Heterogeneity:** To identify geographic clusters of public sentiment and determine how regional economic factors influence these variations using GWR.
3. **Evaluate Diffusion Mechanisms:** To evaluate how different user groups shape the reach and structural persistence of sentiment-laden information within the network.

1.4. Thesis Structure

To provide a clear roadmap of the computational approach, Figure 1.1 illustrates the integrated workflow, spanning from data acquisition and preprocessing to the tripartite analytical core: temporal sentiment trends, spatial heterogeneity via GWR, and retweet network diffusion.

1.5. Literature Review

1.5.1. Societal Impact of International Trade Policies on Public Opinion

International trade policies act as exogenous shocks that trigger widespread socioeconomic anxiety. While early studies focused primarily on macroeconomic indicators such as GDP or trade volume, recent scholarship has transitioned toward capturing the human dimension of these shocks through digital proxies. In terms of platforms, research primarily focused on Twitter to predict market trends or monitor public health events [6, 22].

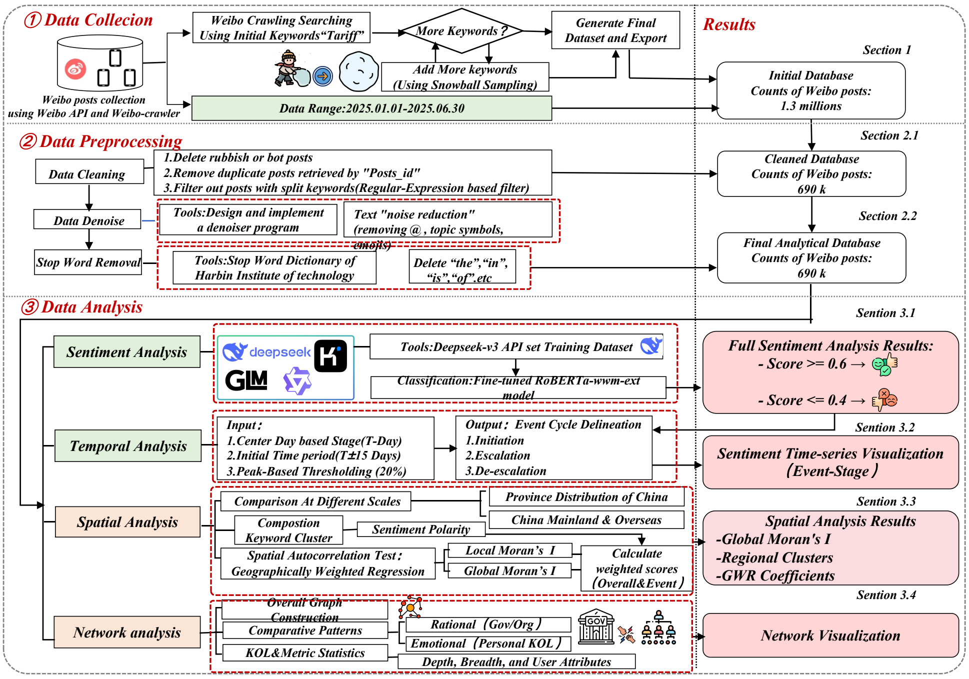

Figure 1.1: The Analytical Framework for Multidimensional Analysisn

Twitter based studies often emphasize stance detection to separate users into opposing camps [5]. However, the discourse ecology on Sina Weibo differs significantly as national policies often show a high degree of nationalist consensus rather than structural polarization [4, 14]. A major limitation of earlier studies is the reliance on stance labels that may collapse into one sided distributions in the Chinese context. Our research overcomes this by focusing on sentiment intensity and emotional volatility, which provides a more nuanced understanding of public mood during the 2025 tariff escalation.

1.5.2. Spatiotemporal Dynamics of Digital Public Opinion

Understanding when and where public sentiment peaks is critical for effective crisis management. Temporally, scholars such as Bollen and Zhu have established that collective mood follows specific decay patterns following external events[6, 28]. AlAziz et al. (2024) [1]further demonstrated that social media reactions often exhibit multi peak structures rather than single bursts. While previous research examined the impact of the COVID 19 pandemic [10, 28], tariff policies differ because they represent an intentional and continuous economic intervention. Spatially, the application of GWR, pioneered by Fotheringham, has allowed researchers to move beyond the assumption that public reaction is uniform across a country. Jasim (2022) and Shi et al. (2022) adopted GWR to map regional differences in pandemic responses[13, 23]. However, few studies have linked these spatial variations directly to specific provincial economic indicators like trade dependency. We address this gap by integrating economic realities with GWR to quantify how regional dependency dictates sentiment polarization.

1.5.3. Mechanisms of Information Diffusion in Social Networks

The impact of a policy shock is determined by how it propagates through social hierarchies. Foundational work by Goel and Vosoughi defined the metrics of breadth, depth, and structural virality [9], proving that different network structures reveal the nature of shared information. Significant research has utilized topology comparison and multi layer network analysis to detect misinformation or false news [19–21]. However, there is less focus on how legitimate policy induced emotions flow across different user strata in a moderated environment. Regarding user influence, scholars debated the influence of follower counts, suggesting that high counts do not always equate to high influence [3, 7]. Our study advances this by evaluating how verified Big V accounts and ordinary users contribute differently to the propagation of sentiment laden cascades, filling the void in research regarding the diffusion paths of official policy discourse.

1.5.4. Sentiment Analysis Methods for Social Media Text

Accurate sentiment detection in Weibo's short and slang heavy text is a prerequisite for reliable research. Traditional machine learning methods often struggle with the linguistic complexity of large scale social media data. The introduction of RoBERTa wwm ext (Wang et al., 2021) revolutionized Chinese NLP by utilizing whole word masking. Recently, using Large Language Models (LLMs) to improve annotation quality has emerged as a new paradigm [2, 11]. While these models are highly accurate, processing millions of posts directly is often too slow or expensive. We solve this by using DeepSeek v3 to generate high quality labels for a subset of data and then fine tuning a RoBERTa model. This approach achieves a balance between high precision and computational efficiency, combining superior reasoning with the deployment scalability required for 1.3 million posts.

1.5.5. Summary of Research Gaps

Despite the extensive literature reviewed above, three critical gaps persist:

1. There is a lack of focus on the multi peak evolution of sentiment in response to

sequential and intentional policy shocks, particularly regarding dynamic event cycle definitions.

2. Research linking provincial level economic dependencies to spatial sentiment non stationarity through GWR in the context of trade wars is insufficient.

3. The structural role of regular users versus institutional accounts in propagating legitimate policy sentiment, as opposed to misinformation, remains under explored.

By addressing these gaps, this thesis provides a novel framework to quantify how society processes major macroeconomic shocks through the integration of LLM technology, cascade analysis, and GWR modeling [12, 15].

Despite the growing body of literature on social media sentiment, three critical gaps remain in understanding public responses to high-stakes geopolitical crises.

First, there is a lack of focus on the continuous evolution of sentiment under sequential policy shocks. Existing studies often focus on sudden, "singular-peak" events or short-lived disasters [1, 22], leaving a gap in quantifying how long-term, sequential policy interventions drive multiple waves of emotional transitions over an extended period [6].

Second, while regional economic exposure drives sentiment differences, most spatial analyses either treat the "Chinese public" as a monolithic entity or focus on micro-scale communities [16, 26]. There is a critical gap in mapping how public anxiety diverges across both export-dependent domestic provinces and international IP locations under a unified global shock [27]. Furthermore, prior work often assumes the relationship between economic drivers and public sentiment is spatially uniform. There is a need for modeling that accounts for spatial non-stationarity to examine how regional economic vulnerability influences collective emotions differently across regions [8, 13].

Third, while text-based sentiment analysis is common, the structural mechanism of how policy-induced emotions spread through social interactions remains under-researched [25]. Although propagation networks are utilized for rumor detection or identifying misinformation [18, 21], there is a lack of research on the hierarchical diffusion mechanism of legitimate policy emotions—specifically, how these emotions flow across different user strata, such as verified "Big V" accounts, enterprises, and ordinary users [3, 7].

While existing studies have extensively analyzed the temporal dynamics of online sentiment, they predominantly employ global regression models (e.g., OLS). These methods assume that the impact of key topics (e.g., Trade, Tech) on sentiment is uniform across the entire nation (spatial stationarity). However, this assumption creates a significant blind spot: it masks the localized nuances where the same topic might trigger opposing emotional responses in different regions due to varying economic structures and cultural backgrounds [8]. Therefore, this study adopts GWR to uncover these hidden spatial structures.

To address these gaps, this study proposes a three-dimensional analytical framework that integrates temporal, spatial, and network topological perspectives. This framework is applied to a large-scale dataset of approximately 1.3 million original Weibo posts collected during the first half of 2025, capturing real-time reactions to macro-policy shifts.

2 | Data and Preprocessing

We collected over 1,300,000 Weibo posts related to 58 tariff-policy keywords. Extracted available metadata includes timestamps, user location, user category (if verified), and IP attributes (when accessible). Text cleaning steps comprised tokenization, stopword removal, deduplication, and noise filtering.

2.1. Data Acquisition

The data acquisition process is divided into two integrated components: the technical implementation of the web crawler and the iterative snowball sampling strategy used to expand the thematic scope of the dataset.

Automated Weibo Crawler

We developed a customized Python based crawler to interface with the Weibo platform and retrieve structured post data. The crawler was configured to capture essential metadata for each post, including the unique user identifier (UID), post content, timestamp, and geotagged IP information. To ensure data integrity and avoid platform rate limiting, the crawler implemented an automated proxy rotation and delayed request mechanism. This technical setup allowed for the continuous collection of raw data throughout the study period, providing the necessary inputs for each round of the keyword expansion process.

Iterative Snowball Sampling

To ensure a comprehensive and representative collection of posts, this study employs a multi stage snowball sampling strategy. Unlike static keyword searches, this method allows the crawling boundary to evolve alongside the public discourse. The process followed three specific phases:

Phase I: Seed Identification The process began with "Tariff" (*Tariff*) as the core seed term. An initial crawl was conducted to retrieve a preliminary set of highly relevant posts, establishing the foundation of the trade conflict discourse.

Phase II: Systematic Keyword Expansion Following each crawling round, we performed a word frequency analysis on the retrieved corpus to identify high frequency co occurring terms and hashtags. By selecting the top ranking keywords and applying domain expertise, we expanded the search parameters. This iterative feedback loop allowed the dataset to grow from generic policy terms to include specific trade actors, affected industries, and macroeconomic outcomes.

Phase III: Final Validation Through three iterative rounds of expansion, a final list of 58 policy specific keywords was curated. This combined approach of frequency based validation and expert judgment ensured the thematic relevance of the final corpus, which comprises approximately 1.3 million original posts spanning from January to June 2025. Approximately 230,000 of these records contain geotagged IP information, enabling spatial analysis at the provincial and country levels.

To provide transparency regarding the technical composition of our findings, we detail the underlying data architecture and the resulting corpus scale.

Table 2.1: Data Schema and Sample Records of the Collected Dataset

Field (ID)	**Description (English)**	**Description (Chinese)**	**Sample (English/Chinese)**
id	Unique post ID	微博唯一标识	5183314549807587
user_nickname	User nickname	用户昵称	日月泽泱
post_time	Post time	发布时间	2025-06-30 19:12
location	Post location	发布位置	NA / 空
repost_count	Repost count	转发数	12
comment_count	Comment count	评论数	9
like_count	Like count	点赞数	4
source_dir	Topic label	话题来源标签	Section 301 investigation / 301调查
ip	IP region	IP归属地	Guangdong / 广东
retweet_id	Retweeted post ID	被转发微博ID	NA / 空
verification	Verification type	认证类型	Gold V / 金V
membership	Membership type	会员类型	Super member / 超级会员
text_raw	Raw text content	原始文本	#特朗普承认美国对中国做过很多事#...

Table 2.2: Overview of the Weibo Dataset

Attribute	**Value**
Collection period	Jan 1 – Jun 30, 2025
Total keywords	58
Total original posts	1,312,847
Total retweet records	40,213
Geolocated posts (IP-based)	230,411
Domestic (China) posts	222,864
Overseas posts	7,547

Table 3.1 outlines the data schema, defining each captured field alongside representative sample records to illustrate the raw information retrieved by the crawler. Following this structural definition, Table 3.2 provides a comprehensive overview of the final Weibo dataset, summarizing the total post volume, unique user counts, and the subset of records available for spatial and network analyses.

2.2. Data Preprocessing

2.2.1. Delete Spam and Bot Posts

Based on typical spam characteristics in the dataset, we classify posts as spam if the post body is empty or extremely short, contains advertising keywords, or contains an excessive number of URLs. After removing obvious spam, we deduplicated posts by their unique ID.

2.2.2. Remove Duplicate Posts Retrieved by Different Keywords

Since the data was collected in batches and subsequently merged, duplicate posts from the same Weibo may exist across multiple crawling sessions.

2.2.3. Filter Out Posts with Split Keywords

We retained `text_clean` as the original clean text and created a new column `tokens` to store tokenization results.

2.2.4. Data Denoising

The denoising pipeline performs the following operations: removing `@user` mentions, converting `#topic#` hashtags to plain text (retaining the inner text), deleting emoji characters, and collapsing redundant whitespace.

Following the automated denoising process, it is essential to validate the distribution and representativeness of the filtered dataset. We conduct a keyword frequency analysis to examine the internal structure of the final corpus and to ensure that the data capture is not disproportionately skewed by a single term. This diagnostic step confirms the breadth of the trade conflict discourse and justifies the use of these specific keywords as the foundation for the subsequent sentiment and network analyses. Figure 2.1 presents the volume of posts and unique users associated with each keyword, ranked by popularity, covering the period from January to June 2025.

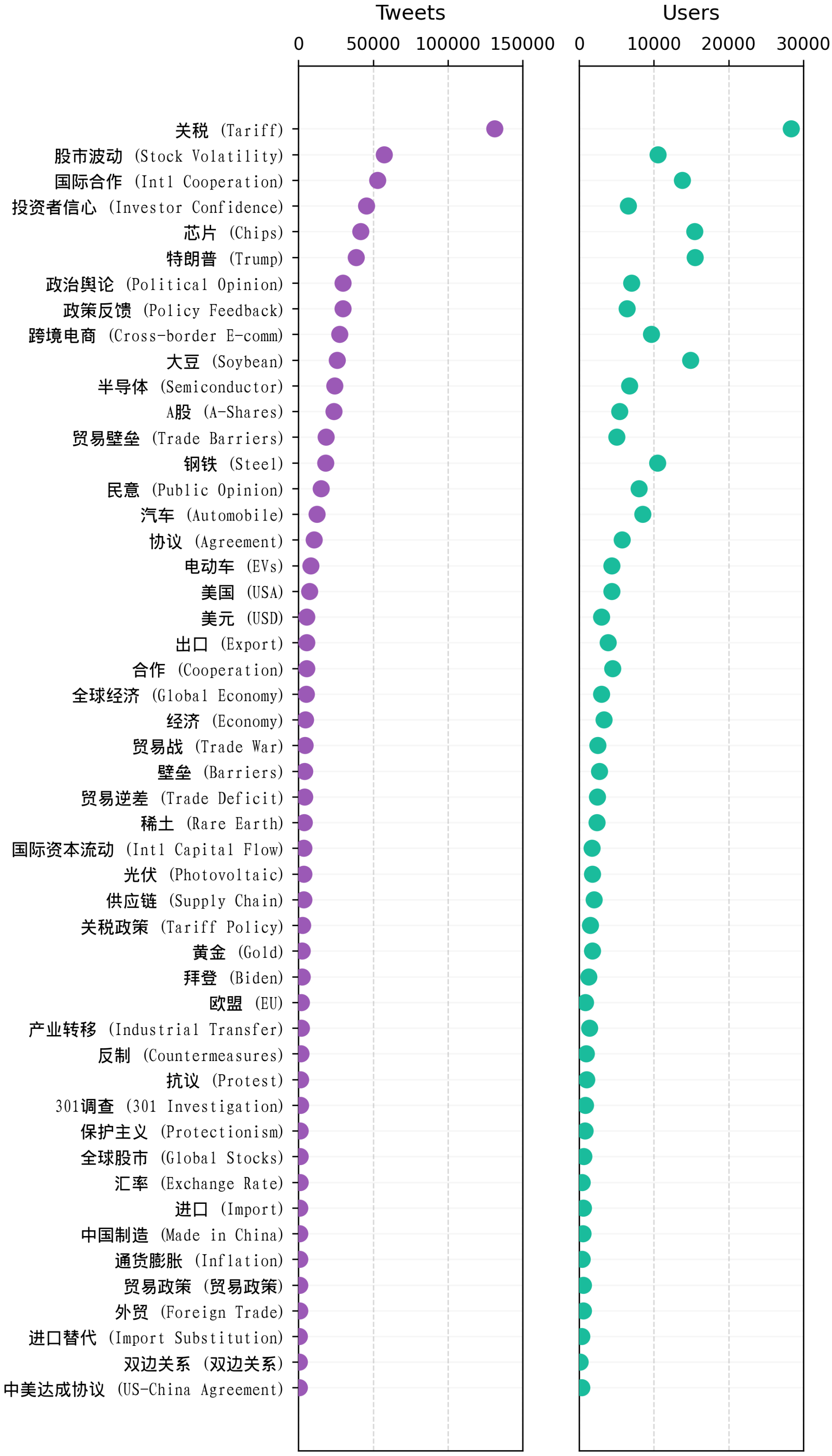

Figure 2.1: Number of posts (purple, left) and users (green, right) captured by each keyword in the final list (ranked by popularity) from January to June, 2025

3 | Methods

This chapter details the multidimensional analytical framework developed to investigate the spatiotemporal sentiment and information diffusion mechanisms of the 2025 trade conflict discourse. The methodology follows a sequential logic, transitioning from micro-level text processing to macro-level spatial and network modelling. The framework initiates with a sentiment analysis pipeline, builds upon these emotional metrics to establish a temporal analysis framework, implements spatial analysis to examine regional disparities, and concludes with retweet network analysis to quantify information flow through cascade metrics.

3.1. Sentiment Analysis

The core of the sentiment classification task utilises a weakly supervised learning paradigm to process 1.3 million records. Given the limitations of manual labelling at this scale, we implement a knowledge distillation pipeline where the DeepSeek-v3 Large Language Model acts as a teacher to generate high-precision silver labels for a balanced training set. These labels are then used to fine-tune a specialised RoBERTa-based classifier (Chinese-RoBERTa-wwm-ext), optimising the model for the specific linguistic nuances of Chinese policy discourse on Weibo.

3.1.1. Sentiment Labeling and Fine-tuning Pipeline

We first randomly sampled 500 complete Weibo posts and applied the same prompt template to obtain three-class sentiment predictions from four mainstream LLMs: DeepSeek-V3, GLM-4.6, Qwen-Turbo, and Tongyi. We then randomly selected an additional 50 posts for manual sentiment annotation and compared model predictions against human labels to compute accuracy; DeepSeek achieved the highest accuracy (80.00%) and was

Algorithm 3.1 Weakly Supervised Sentiment Labeling and Fine-tuning Pipeline

Require: Raw Weibo dataset $\mathcal{D}_{raw}$, Pre-trained RoBERTa model M_{base}, LLM Teacher (DeepSeek-v3)
Ensure: Fine-tuned Sentiment Classifier M_{tuned}

```
Sample a balanced subset D_sub ⊂ D_raw across different keywords;
{Knowledge Distillation Phase}
for each post x_i ∈ D_sub do
    Prompt DeepSeek-v3 to generate a pseudo-label y_i ∈ {−1, 0, 1};
    Store pair (x_i, y_i) into silver training set D_silver;
end for
{Model Fine-tuning Phase}
Initialize M_tuned with weights of M_base;
while not converged do
    Train M_tuned on D_silver using Cross-Entropy Loss: L = −∑ y log(p);
    Update weights via AdamW optimizer with LR = 2 × 10^−5;
end while
return M_tuned
```

therefore selected as the labeler for the training data. Next, we sampled 2,000 posts as a fine-tuning set and used DeepSeek to assign negative/neutral/positive labels. These labeled samples were used to fine-tune a Chinese RoBERTa-wwm-ext model for three-class sentiment classification with a BERT sequence classification head (single-label setting).

The reason we chose the pre-trained RoBERTa-wwm-ext model [24] for sentiment classification is based on empirical studies on the Weibo dataset that validated its superiority. Wang et al. (2021) compared BERT, RBT3, and RoBERTa-wwm-ext on a dataset containing over 100,000 Weibo comments and found that RoBERTa-wwm-ext achieved state-of-the-art performance (accuracy: 80.89%, macro F1 score: 68.06%).Table 3.1 presents the parameter information for the RoBERTa Fine-tuning Pipeline.

Table 3.1: Hyperparameters for RoBERTa Fine-tuning Pipeline

Hyperparameter	Value
Pre-trained Model	Chinese-RoBERTa-wwm-ext
Teacher Model (Labels)	DeepSeek-v3 (LLM)
Learning Rate	2×10^{-5}
Batch Size	32
Max Sequence Length	128
Epochs	5
Optimizer	AdamW
Loss Function	Cross-Entropy

Training ran for 3 epochs with 339 total steps and a batch size of 16. The best validation checkpoint appeared at epoch 2 (step 226) with a validation loss of 0.7551, followed by a slight increase at epoch 3. The final fine-tuned model and tokenizer were saved, and the full training process and loss curve were recorded. Finally, the fine-tuned RoBERTa-wwm-ext model was applied to the full Weibo corpus to generate sentiment labels for downstream diffusion analysis.

To identify sentiment trends, we aggregate individual sentiment scores into daily mean trajectories. We apply a moving average smoothing technique to filter daily noise and highlight persistent emotional shifts following policy announcements. The temporal analysis focuses on measuring the decay rate of emotional peaks and the correlation between event milestones and sentiment volatility.

3.2. Temporal Analysis

To identify sentiment trends, we aggregate individual sentiment scores into daily mean trajectories and apply a 7-day moving average smoothing technique to filter daily noise and highlight persistent emotional shifts following policy announcements. The temporal analysis focuses on measuring the decay rate of emotional peaks and the correlation between event milestones and sentiment volatility.

3.2.1. Dynamic Event Phase Segmentation

Event cycles are delineated dynamically rather than fixed by calendar dates. Given a daily post volume sequence, we first apply 7-day Simple Moving Average (SMA) smoothing. A significance threshold H is then set at 20% of the global peak volume. An active phase begins when the smoothed volume crosses H from below, and ends when it falls back below H. The local maximum within each active phase is designated as the phase centre day (T-Day). A symmetric ±15-day window around T-Day defines the analytical event window for sentiment composition analysis. This approach identifies three distinct policy lifecycle phases: Policy Initiation (centred 2025-02-13), Conflict Escalation (centred 2025-04-09), and De-escalation (centred 2025-05-14), corresponding to the major tariff announcement events of the period.

Table 3.2: Definition of Policy Lifecycle Phases and Event Windows

Phase Name	Start Date	End Date	Central Day
Policy Initiation	2025-02-06	2025-02-27	02.13
Conflict Escalation	2025-04-02	2025-04-23	04.09
De escalation Period	2025-05-07	2025-05-28	05.14

Algorithm 3.2 Dynamic Event Phase Segmentation

Require: Daily post volume sequence $V = \{v_1, v_2, ..., v_n\}$, Threshold $\tau = 0.2$
Ensure: Event Phase Boundaries $\{T_{start}, T_{peak}, T_{end}\}$
Apply Moving Average smoothing: $V_{smooth} = \text{SMA}(V, \text{window} = 7)$;
Identify global peak intensity $V_{max} = \max(V_{smooth})$;
Define significance threshold $H = V_{max} \times \tau$;
for each day t in V_{smooth} **do**
  if $v_t > H$ and $v_{t-1} \leq H$ **then**
    Mark $T_{start} = t$ {Start of an active phase}
  else if $v_t < H$ and $v_{t-1} \geq H$ **then**
    Mark $T_{end} = t$ {Return to stabilization}
  end if
end for
return Registered phases with local maxima as T_{peak}

3.3. Spatial Analysis

The spatial dimension evaluates how geographic and economic factors influence public sentiment. After testing for spatial autocorrelation using Global Moran's I, we implement GWR to model the relationship between sentiment and regional variables. This local regression approach allows the parameters to vary across space, capturing how the impact of tariff policies differs between export-dependent coastal provinces and inland regions.

3.3.1. Spatial Autocorrelation

Global Moran's I is computed to assess whether provincial-level sentiment scores exhibit statistically significant spatial clustering or dispersion relative to a random pattern. A positive and significant I indicates that provinces with similar sentiment tend to cluster geographically, justifying a local regression approach. Local Moran's I (LISA) is subsequently used to identify specific clusters of high-high or low-low provincial sentiment, as well as spatial outliers. The Getis-Ord G_i^* statistic is used to identify statistically significant spatial hot spots and cold spots:

The local spatial association for a given unit i is calculated as:

$$I_i = \frac{(x_i - \bar{X})}{S^2} \sum_{j=1, j \neq i}^{n} w_{ij}(x_j - \bar{X}) \tag{3.1}$$

where w_{ij} is the spatial weight between units i and j, x_j is the attribute value at j, $\bar{X}$ is the global mean, S is the global standard deviation, and n is the total number of spatial units.

3.3.2. Geographically Weighted Regression Model Specification

GWR is employed to capture the spatial non stationarity of the relationship between thematic discourse composition and provincial level sentiment. Rooted in Tobler's First Law of Geography—which posits that near things are more related than distant things—GWR allows the influence of specific topics to vary across space rather than assuming a universal effect.

Unlike OLS, which produces a single global parameter estimate, GWR calibrates a separate local model at each spatial unit i by down weighting observations that are geographically distant from the focal point. The model is specified as:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{3.2}$$

where y_i is the mean sentiment score for province i; (u_i, v_i) are the geographic coordinates of the centroid of province i; $\beta_k(u_i, v_i)$ represents the locally estimated coefficient for thematic predictor k at location i; and ε_i is the residual error term. The independent variables x_{ik} represent the normalized post volume proportions for four keyword clusters: *Strategic*, *Trade*, *Livelihood*, and *Technology*.

To address the uneven density of Weibo users—denser in coastal economic hubs and sparser in the interior—we utilize an adaptive bi square kernel. The bandwidth (BW) is optimized by minimizing the corrected Akaike Information Criterion (AICc), ensuring a balance between model complexity and goodness of fit. Furthermore, to ensure the validity of the local estimates, we monitor the Condition Number to detect potential local multicollinearity, ensuring that the thematic predictors are not highly correlated within the local neighborhoods.

Model performance is evaluated by comparing the local R^2 surface against the global OLS benchmark. This allows us to identify where the selected thematic topics provide the strongest explanation for sentiment variation. Finally, we apply Global Moran's I to the GWR residuals to confirm that the model has successfully absorbed the spatial dependence inherent in the raw data. All computations are executed using the `mgwr` Python library.

3.4. Retweet Network Analysis

The final component analyzes the structural mechanisms of information diffusion. We construct retweet networks where nodes represent users and edges represent information flow. To quantify the diffusion patterns, we calculate cascade-level metrics including Breadth (total size), Depth (number of layers), and Structural Virality. These metrics allow for a comparative analysis of how different user categories, such as institutional accounts and individual influencers, contribute to the persistence of sentiment cascades.

We construct the retweet network where nodes represent users and edges represent retweet directions, weighted by frequency [17, 18].

3.4.1. Network Construction and User Typology

The retweet network is constructed from the 40,213 retweet records in the dataset. Each directed edge (u → v) represents a retweet event in which user u shared content originating from user v. Unique chains of retweet relationships emanating from a single root post are modelled as rooted trees, forming information cascades. This tree-based representation captures hierarchical propagation and enables decomposition of diffusion into breadth (horizontal reach) and depth (vertical penetration) dimensions. Users are classified into five strata based on Weibo's authentication system: (1) Gold V — individual celebrities and verified opinion leaders; (2) Red V — government agencies and official administrative accounts; (3) Yellow V — enterprise and commercial accounts; (4) Blue V — official institutional media organisations; and (5) Regular — unverified users. This typology enables comparative assessment of how institutional authority, commercial identity, and individual influence differentially shape propagation architecture. To examine whether cascade topology differs between factual policy content and emotionally reactive content, cascades are additionally classified as Mainstream (Rational) — originating from Blue V or Red V accounts with neutral-to-moderate sentiment — or Emotionally Charged — high-polarity cascades (score > 0.75 or < 0.25) from non-institutional accounts. This binary typology is used in the structural comparison presented in Section 4.4.1.

3.4.2. Impact Threshold Rationale

Raw cascade statistics are dominated by a large number of trivial single-retweet events that inflate mean values and obscure meaningful propagation patterns. To isolate substantively significant diffusion events, we apply two size thresholds: a Basic threshold (Size > 2), which excludes isolated single-retweet dyads while retaining all minimally networked cascades; and an Impact threshold (Size > 5), which restricts analysis to cascades that have demonstrated meaningful social traction. Comparing results across both thresholds reveals the 'super-broadcaster' activation effect — the disproportionate amplification that certain user strata achieve once content crosses a minimum engagement threshold — which would be obscured if only a single threshold were applied.

3.4.3. Cascade Metrics

For each cascade subgraph $G = (V, E)$, the breadth is defined as the number of nodes:

Cascade Breadth (Size). The total number of nodes in the cascade, measuring aggregate reach:

$$\text{Size} = |V|. \tag{3.3}$$

Cascade Depth (Tree Depth). The maximum shortest-path distance from the root r to any node v, quantifying the number of propagation layers:

$$\text{Depth} = \max_{v \in V} d(r, v). \tag{3.4}$$

Cascade Diameter (Graph Diameter). The maximum pairwise shortest-path distance between any two nodes in the undirected cascade subgraph, measuring topological span:

$$\text{Diameter} = \max_{u,v \in V} d(u, v). \tag{3.5}$$

CCDF (Tail Probability). For cascade size X, the Complementary Cumulative Distribution Function quantifies the probability that a cascade exceeds a given size threshold x:

$$\text{CCDF}(x) = P(X > x) = 1 - F(x). \tag{3.6}$$

The CCDF is plotted on a log-log scale to reveal power-law behaviour and heavy-tailed distributions. A higher tail for a given sentiment class indicates that posts of that polarity are systematically more likely to generate large or deep cascades, reflecting asymmetric virality dynamics. To differentiate broadcast-type diffusion (shallow, wide) from chain-type diffusion (narrow, deep), average depth per cascade is additionally reported as a proxy for the depth-to-breadth ratio.

4 | Results and Discussion

4.1. Sentiment Analysis Results

This section presents the temporal evolution and structural composition of public sentiment on Weibo. By examining the daily fluctuations in posting volume and emotional intensity, we establish the macroeconomic baseline for the 2025 tariff event. The analysis focuses on identifying how sentiment reacts to specific policy shocks and how these responses stabilize over the event lifecycle.

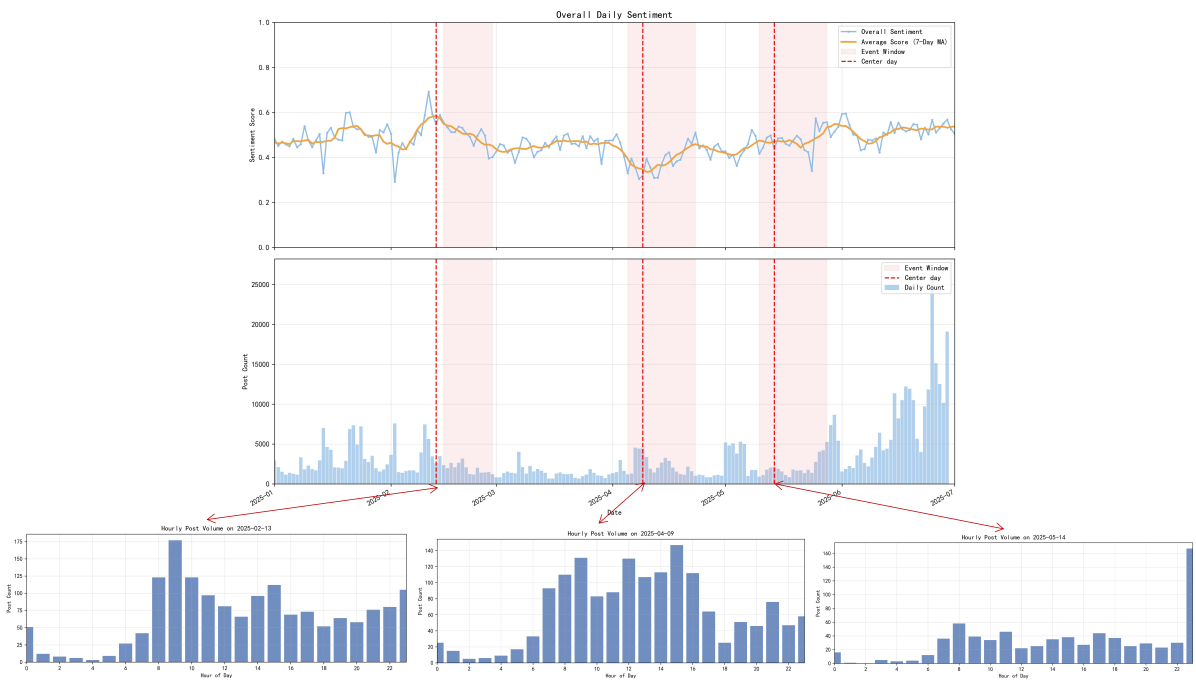

Figure 4.1: Temporal Trends of Overall Sentiment and Daily Posting Volume with Event Center Days

Figure 4.1 illustrates the daily sentiment trajectories alongside posting volumes, revealing a direct synchronization between policy announcements and public attention peaks. We observe that sentiment intensity is highly volatile during the initial intervention period, with sharp fluctuations corresponding to the introduction of tariff measures. This visualization allows us to identify the critical event windows where public discourse transitions from immediate shock to sustained discussion.

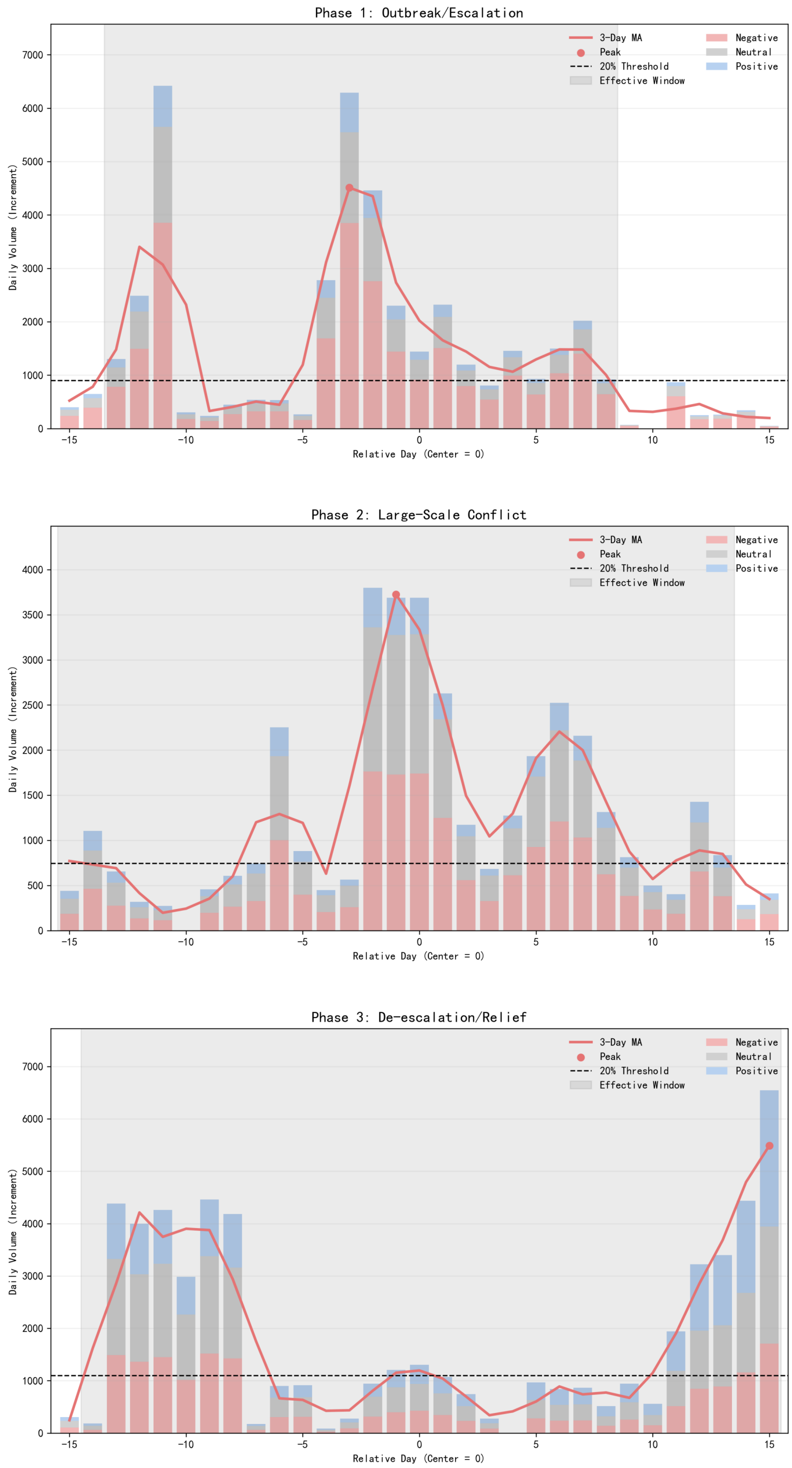

Figure 4.2: Event Lifecycle by Phase: Sentiment Composition and Volume Increment

To further quantify these shifts, Figure 4.2 details the event lifecycle through different phases. By utilizing our dynamic detection method, we observe the sentiment composition and volume increment across the policy lifecycle. This breakdown reveals that while volume peaks during the conflict escalation phase, the sentiment composition shifts as the public moves through the shock, escalation, and de escalation stages. This temporal foundation provides the necessary context for the subsequent global comparison and the province level spatial analysis.

4.2. Global Comparison Results

Following the sentiment analysis of the primary domestic dataset, we situate the Chinese public response within the broader international discourse by comparing domestic Weibo data with overseas Twitter (X) data. The analysis proceeds in three steps: we first examine the temporal alignment of engagement volumes between China and the overseas aggregate; we then analyse the divergence in daily sentiment trajectories; and we zoom in on the China-US. dyad as the two primary actors in the trade conflict. The section concludes by validating the geographic representativeness of the overseas sample and by mapping the domestic provincial distribution of engagement as a bridge to the subsequent spatial analysis.

4.2.1. Engagement Volume: China vs. Overseas

We begin by examining the scale and temporal alignment of the discussion. Figure 4.3 plots absolute daily post volumes for the two cohorts across the full study period. The domestic corpus consistently dwarfs the overseas sample in raw count, reflecting both the size of Weibo's user base and the direct salience of U.S. tariff policy for Chinese citizens. Despite this scale difference, both series share identifiable peaks centred on the three major policy events, confirming that the same external shocks register as attention triggers across geographies.

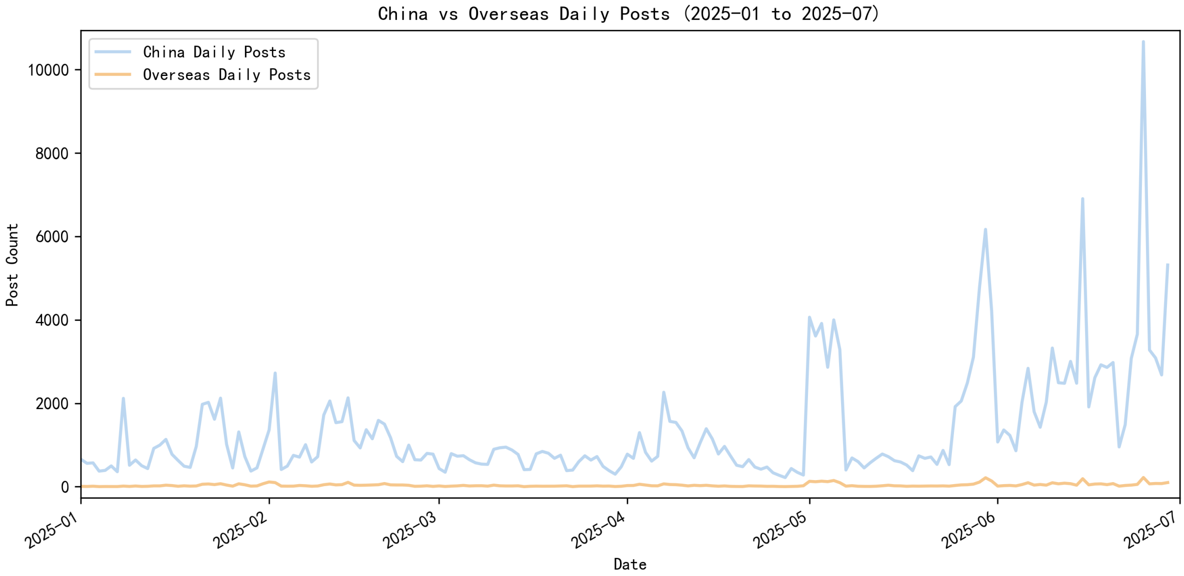

Figure 4.3: Daily Tweet Volume: China vs Overseas

Because absolute counts obscure relative response intensity, Figure 4.4 normalises each series to its own maximum. The normalised view reveals that overseas engagement exhibits a *pulse-like*sharply pattern, narrow peaks followed by rapid decay where the domestic response follows a *progressive* trajectory, featuring a gradual pre-event build-up and a sustained post event tail. From May 2025 onward, the domestic baseline rises steadily, signalling a structural transition from event-driven spikes to sustained policy discourse, while the overseas baseline remains volatile and low between events.

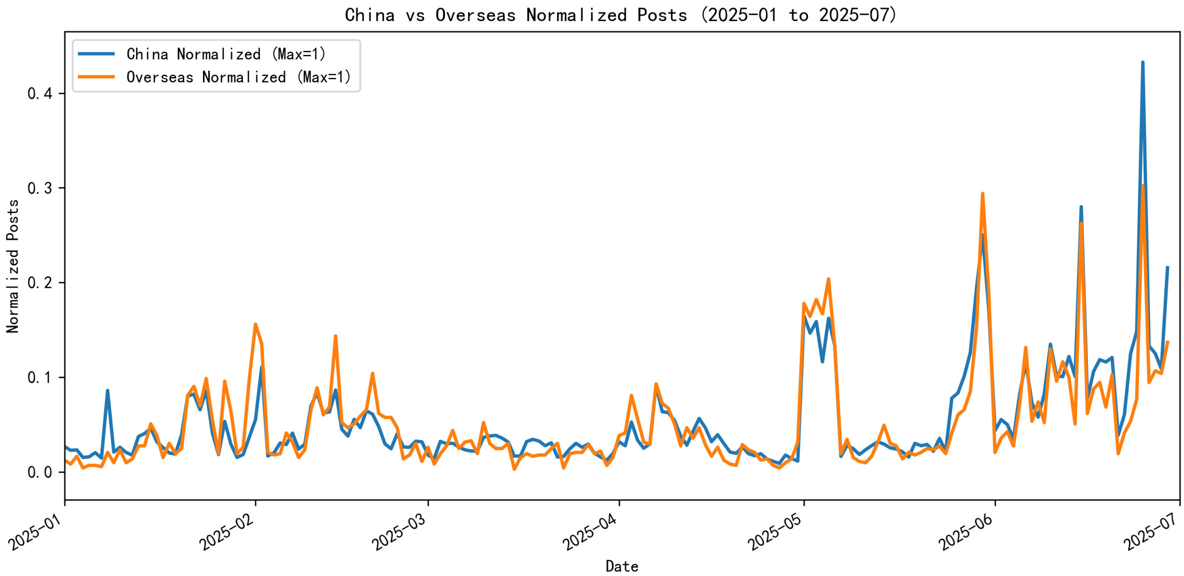

Figure 4.4: Normalized Daily Tweet Volume: China vs Overseas

4.2.2. Sentiment Trajectories: China vs. Overseas

Beyond volume, Figure 4.5 tracks daily mean sentiment scores for each cohort, revealing fundamental differences in emotional stability. Domestic sentiment oscillates within a narrow band of approximately [0.45, 0.60] throughout the study period, displaying a buffered response in which even major policy announcements produce gradual, smooth shifts rather than abrupt reversals. Overseas sentiment, by contrast, exhibits substantially higher amplitude and wider variance, characterised by a peak-and-trough structure in which scores surge or drop sharply in response to individual news cycles. This greater susceptibility to short-term narratives indicates a higher polarisation risk for overseas audiences, while the domestic trajectory reflects the stronger narrative continuity and emotional equilibrium consistent with a more consolidated media environment.

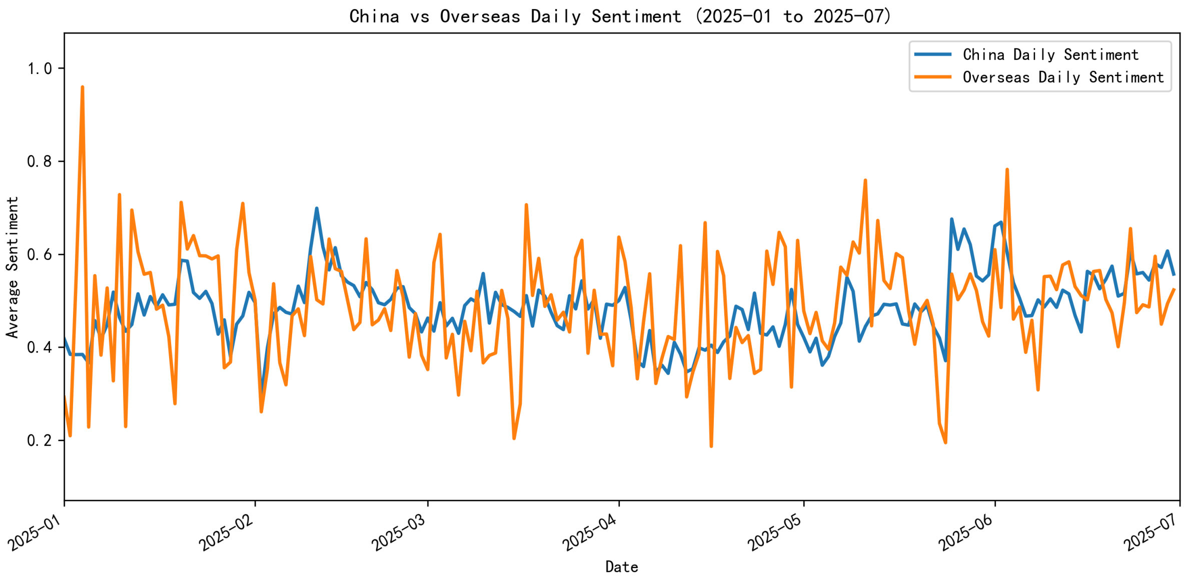

Figure 4.5: Daily Sentiment Trajectories: China vs Overseas

4.2.3. The China–U.S. Dyad

Because the United States accounts for approximately 30.4% of the overseas sample and serves as the policy origin, we introduce a targeted China–U.S. comparison.Figure 4.6 plots normalised daily volumes for the two countries. U.S. engagement is defined by *short-burst* dynamics — rapid escalation immediately following a policy announcement, followed by near-complete dissipation within days — contrasting sharply with the domestic sustained-engagement model. This burstiness suggests that U.S. discourse is organised around individual events rather than evolving as a continuous deliberative process.

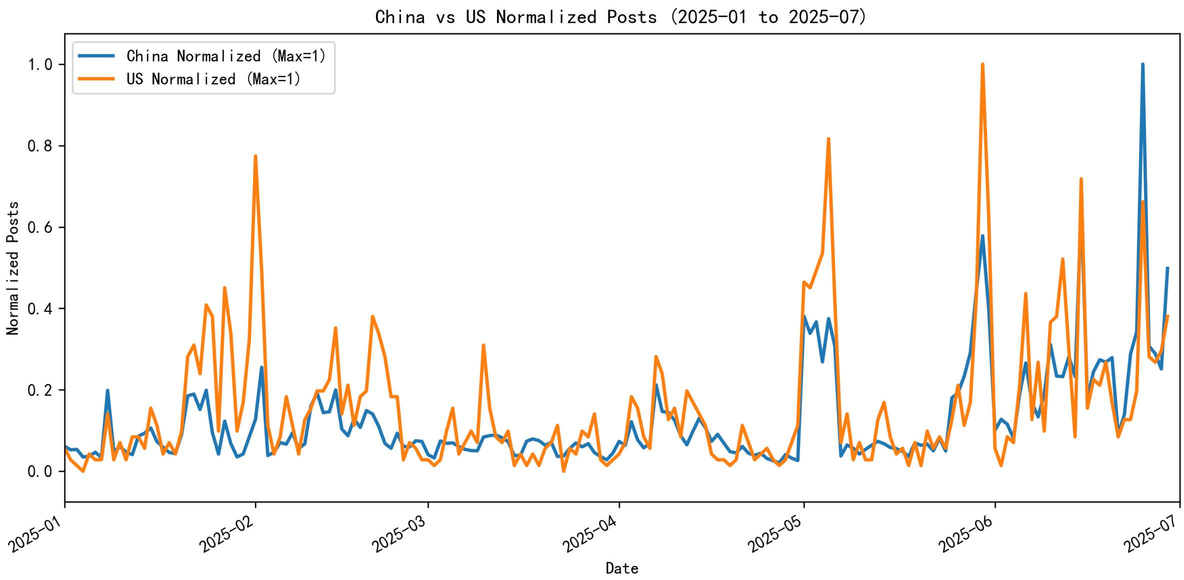

Figure 4.6: Normalized Daily Tweet Volume: China vs US

Figure 4.7 confirms that the emotional contrast is equally stark.The U.S. sentiment curve exhibits the most pronounced spikes and highest emotional tension of any sub-group examined, with amplitude significantly exceeding both the domestic trajectory and the aggregate overseas average, identifying U.S. users as the primary driver of overseas emotional heterogeneity.

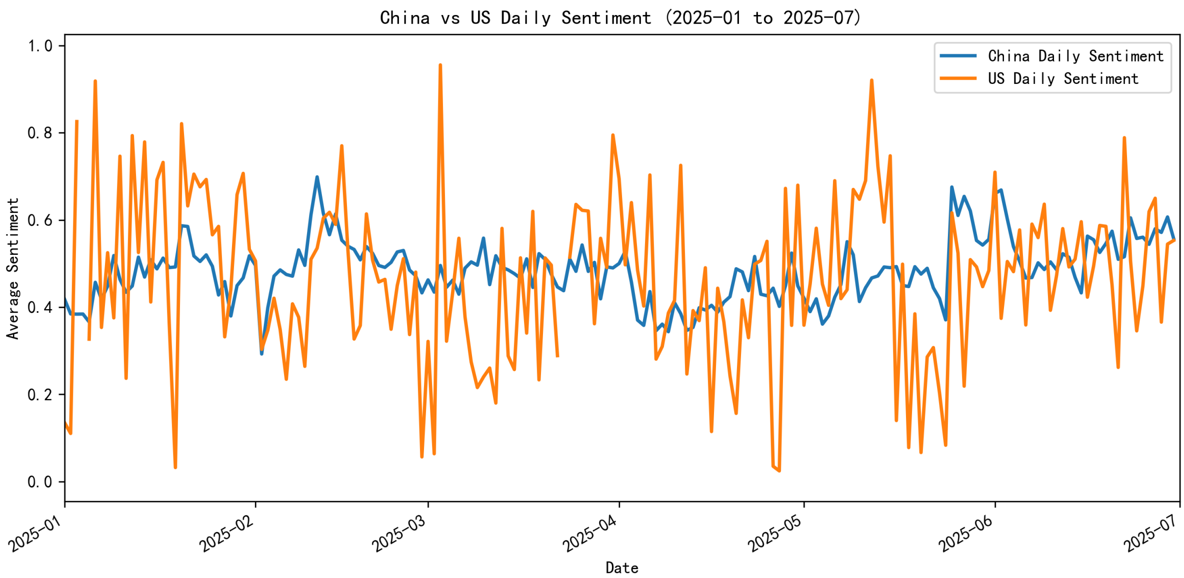

Figure 4.7: Daily Sentiment Trajectories: China vs US

4.2.4. Geographic Distribution of the Overseas Sample

To validate that the overseas sentiment baseline is not disproportionately skewed by a single region beyond the United States, we decompose the overseas sample by country. Figure 4.8 maps absolute post counts, confirming that the United States, followed by English-speaking and East Asian countries, contributes the largest share of overseas discourse.

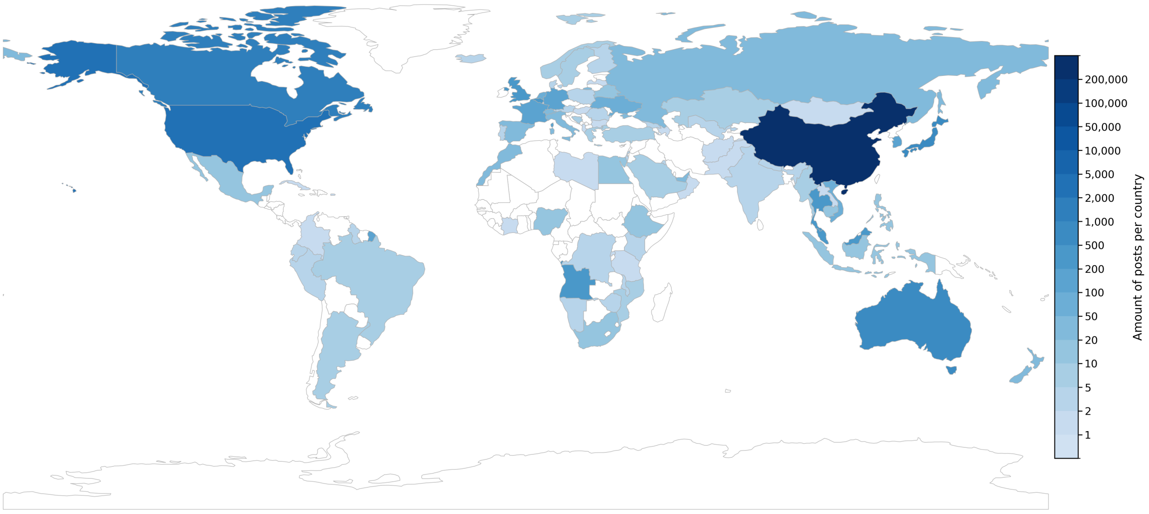

Figure 4.8: Total Posts by Overseas Country

Figure 4.9 normalises by national population, revealing that each capita engagement is highest in countries with large Chinese diaspora communities and close trade relationships with both the U.S. and China. This each capita view confirms that the overseas sample, while U.S.heavy in absolute terms, represents a geographically diverse international discourse rather than a single country perspective.

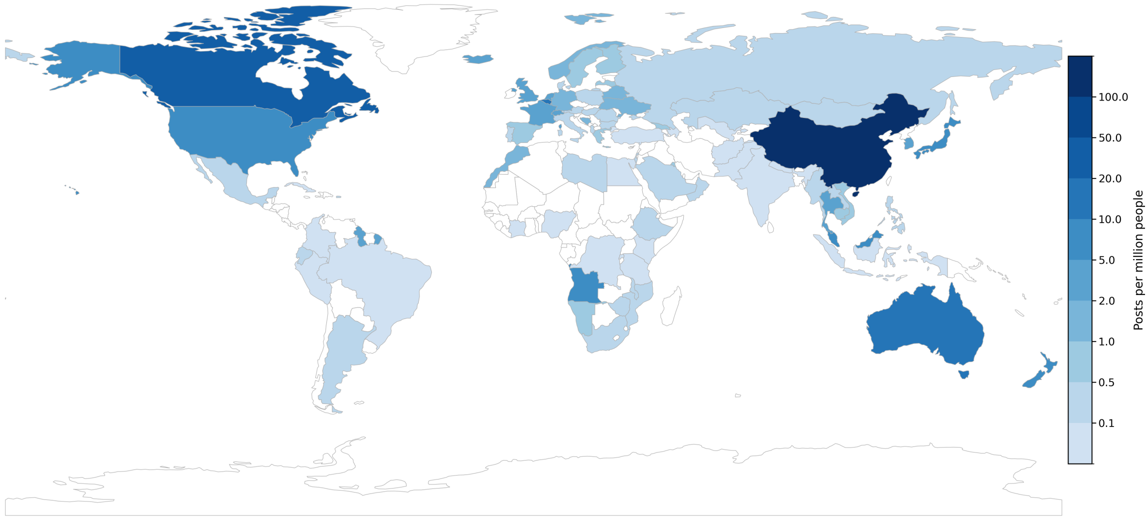

Figure 4.9: Posts per Million by Overseas Country

4.2.5. Geographic Distribution of Weibo Engagement

Before examining the spatial non-stationarity of sentiment drivers, we analyse the geographic distribution of the geolocated domestic dataset across China's provinces. This descriptive stage is essential for identifying regional hotspots of public engagement and for contextualising the subsequent GWR results within China's uneven economic geography.

Figure 4.10 maps absolute post volume by province. Engagement is heavily concentrated in the eastern coastal belt, with Guangdong,Jiangsu, and Zhejiang generating the highest raw counts,a pattern that reflects both population density and the concentration of export-oriented manufacturing in these provinces.

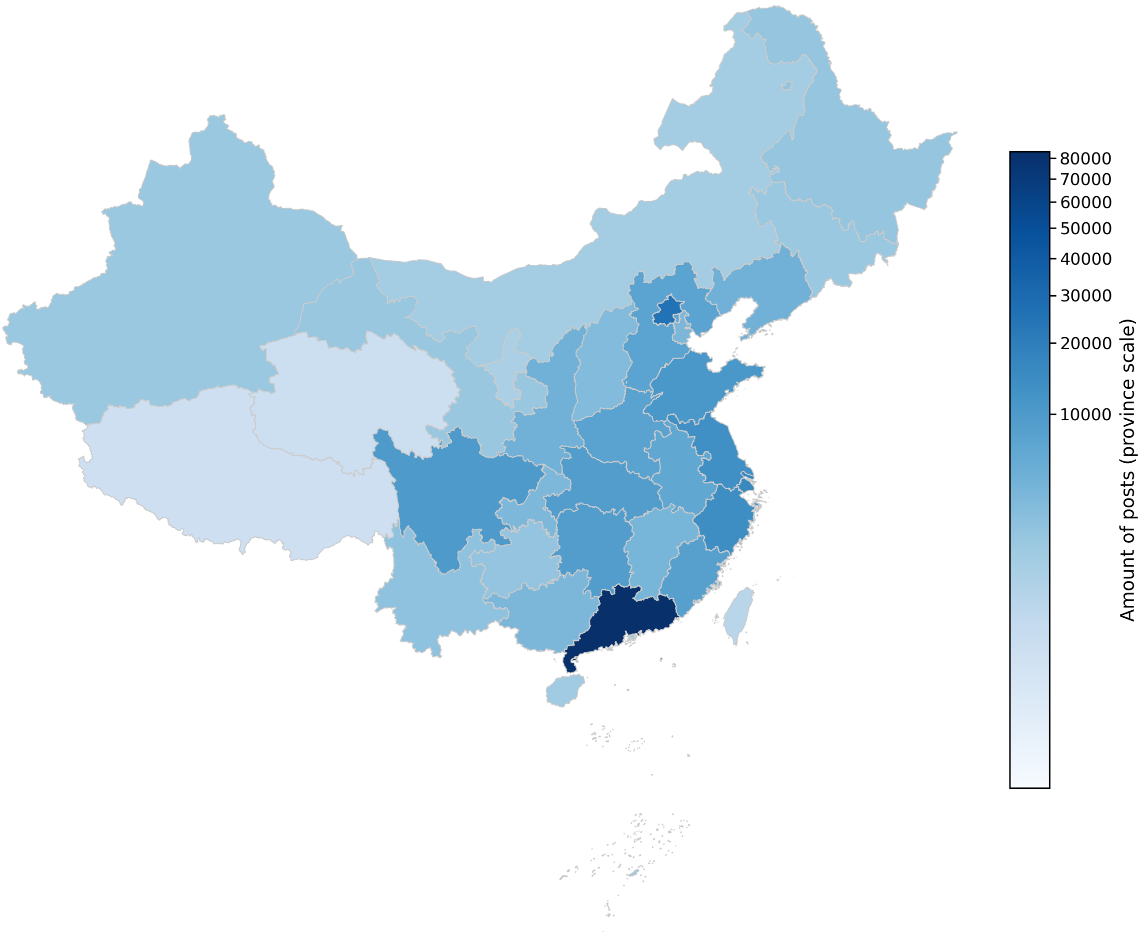

Figure 4.10: Spatial Distribution of Total Weibo Post Volume Across Chinese Provinces

However, raw volume conflates population size with genuine policy sensitivity. Figure 4.11 addresses this by normalising post counts per million residents. Under this normalisation, municipalities such as Beijing and Shanghai maintain high per-capita engagement, reflecting the concentration of politically attentive and economically informed urban populations. Conversely, some populous coastal provinces see their apparent dominance attenuated, indicating that their high raw volumes are partly a function of population size rather than disproportionate policy concern.

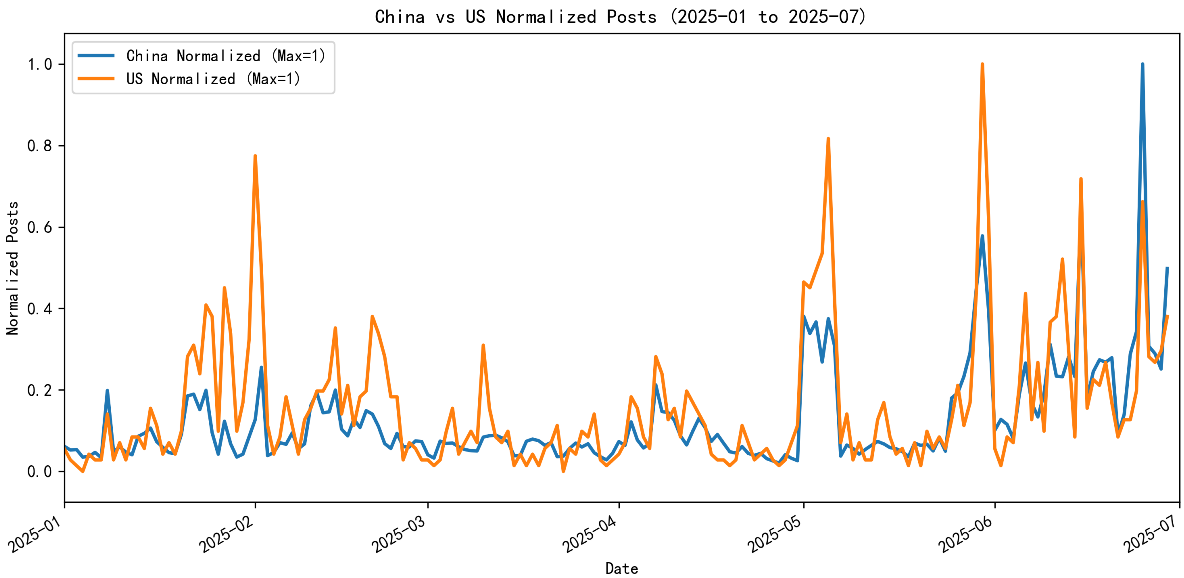

Figure 4.11: Comparative Temporal Trajectory of Normalized Daily Post Volume Between China and the United States

The spatial distribution of discourse reveals a significant concentration in eastern coastal regions and major economic hubs. While absolute post volume (Figure 4.10) is naturally higher in populous provinces such as Guangdong, Jiangsu, and Zhejiang, the population normalized data (Figure 4.11) provides a more nuanced identification of regional sensitivity. The high intensity engagement observed in these trade oriented provinces suggests a direct correlation between regional economic exposure and public attention. This geographically structured variation confirms that tariff related discourse is not uniform across space, thereby providing the empirical justification for the GWR analysis in the following section, where we model the spatial non stationarity of sentiment drivers.

4.3. Spatial Analysis Results

This section presents the empirical results of the geographic analysis, focusing on the regional distribution and thematic drivers of public sentiment across China. We follow a sequential analytical framework: first, establishing a descriptive baseline of keyword clusters; second, validating the superior performance of the GWR model over the global benchmark; and finally, analyzing the spatiotemporal evolution of regression coefficients. This integrated approach reveals the profound connection between regional economic structures and digital public discourse.

4.3.1. Thematic Baseline and Sentiment Polarity

To understand the semantic foundation of the discourse, we first categorize the domestic data into four functional clusters. Table 4.1 provides the descriptive statistics for these clusters, highlighting the frequency and mean sentiment of the most influential keywords.

Table 4.1: Descriptive Statistics of Keyword Clusters and Sentiment Mean

Cluster	Core Keywords (Examples)	Freq.	Mean Sent.	Total Posts
Strategic	USA (美国)	186,019	0.46	341,840
	Trump (特朗普)	137,777	0.49	
	Cooperation (合作)	100,461	0.53	
	Dollar (美元)	71,606	0.53	
	EU (欧盟)	32,394	0.59	
Trade	Tariffs (关税)	193,876	0.52	264,389
	Export (出口)	54,292	0.49	
	Import (进口)	41,672	0.50	
	Tariff Policy (关税政策)	33,779	0.50	
	Supply Chain (供应链)	26,615	0.41	
Livelihood	Economy (经济)	132,519	0.48	238,094
	A-Shares (A股)	69,851	0.48	
	Gold (黄金)	34,385	0.67	
	Public Opinion (民意)	17,057	0.32	
	Exchange Rate (汇率)	11,160	0.58	
Tech	Automobile (汽车)	60,930	0.50	196,348
	Chips (芯片)	62,843	0.48	
	Semiconductors (半导体)	42,730	0.58	
	New Energy (新能源)	25,250	0.49	
	Rare Earth (稀土)	13,839	0.57	

Note: Only the top 5 keywords by frequency are displayed for each cluster. Sentiment scores are normalized between 0 (negative) and 1 (positive).

The emotional distribution of these clusters is further visualized in Figure 4.12, which compares the sentiment polarity across the four dimensions. This descriptive foundation confirms that while Strategic and Trade topics dominate in volume, Livelihood and Technology topics exhibit higher emotional volatility.

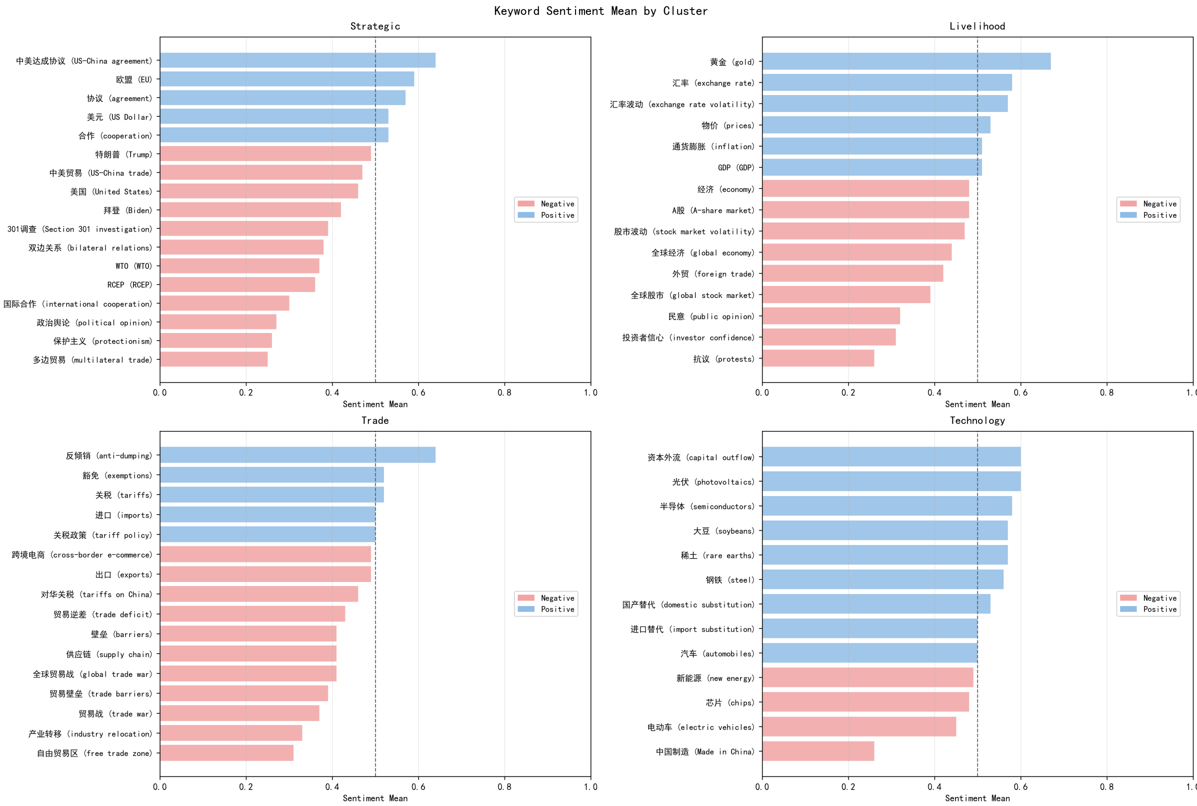

Figure 4.12: Comparative analysis of sentiment polarity: Strategic, Trade, Livelihood, Technology

4.3.2. Model Validation and Local Fitting Intensity

Moving from descriptive statistics to spatial modeling, we evaluate the effectiveness of the GWR framework. As shown in Table 4.2, the GWR model significantly outperforms the OLS benchmark, with the R^2 increasing from 0.0306 to 0.6004. This dramatic improvement indicates that the spatial model effectively captures local variations that a global regression fails to address.

Table 4.2: Model Fit and Spatial Autocorrelation (Overall vs. Conflict Escalation)

Stage	**Global Moran's I**	**p value**	**OLS R^2**	**GWR R^2**	**AICc**	**BW**
Overall	-0.0810	0.6714	0.0306	0.6004	-0.58	27.0
Conflict escalation	-0.1705	0.0880	0.2570	0.6086	88.44	9.0

To further verify the spatial disparities in model performance, Table 4.3 provides the descriptive statistics for the local R^2 values. The results confirm that the fitting intensity is not uniform across provinces, with R^2 values ranging from 0.4192 to 0.6663, thus justifying a localized approach.

Table 4.3: Descriptive Statistics of Local R^2 (Overall)

N	Mean	Median	Min	Max	Std
30	0.5601	0.5650	0.4192	0.6663	0.0731

Figure 4.13 utilizes these values to visualize the spatial disparities in fitting intensity. By mapping the local R^2, we provide a geographic representation of the model's varying explanatory power, allowing us to examine the non stationarity of the relationship between public sentiment and thematic structures.

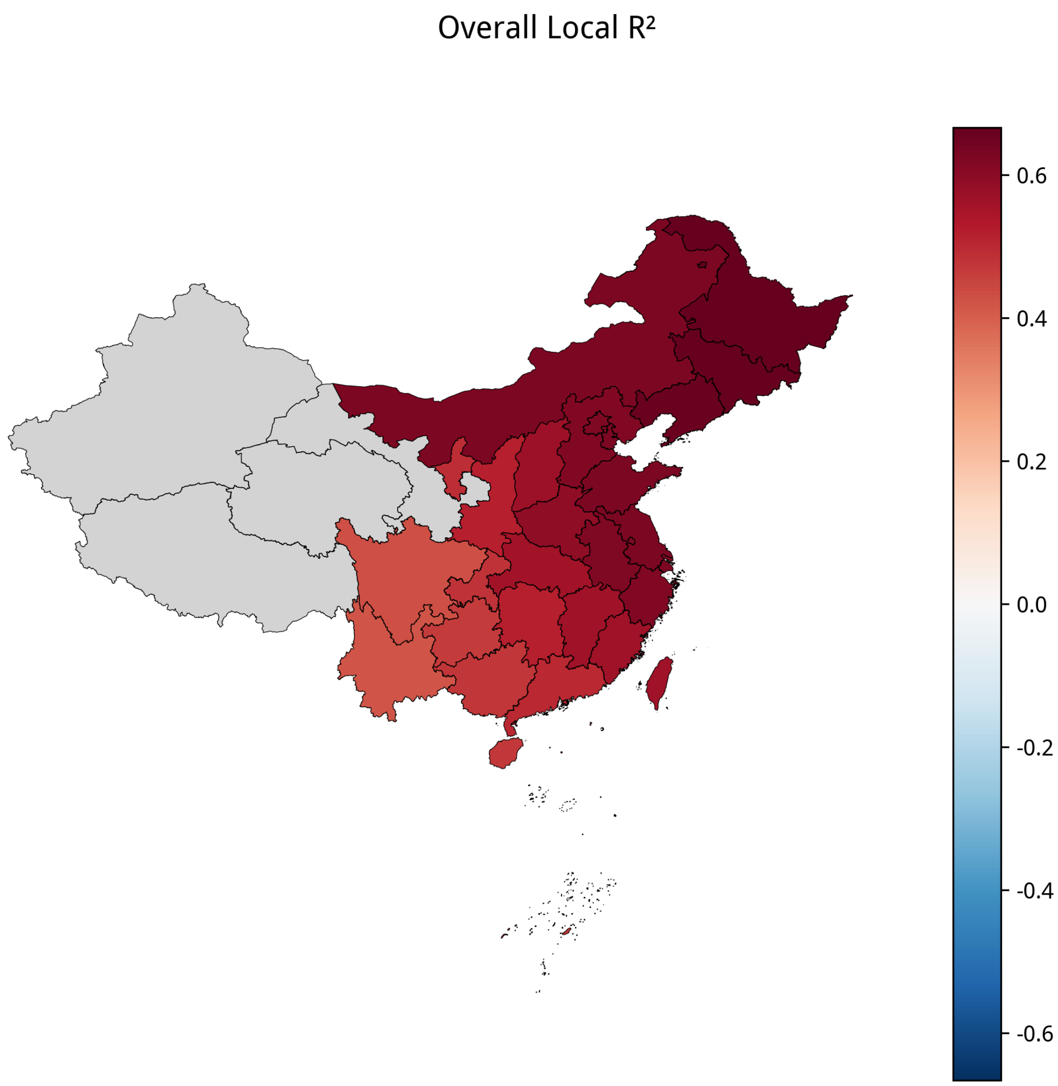

Figure 4.13: Spatial distribution of local R^2 values

The map in Figure 4.13 clearly illustrates that provinces with darker colors, primarily in the eastern and northeastern regions, possess higher R^2 values. This suggests that sentiment dynamics in these areas are explained more effectively by the model. In contrast, the lighter shading in western regions indicates a relatively weaker explanatory power. This pattern demonstrates the significant spatial heterogeneity of the sentiment driven mechanism, as the model performs significantly better in trade and technology intensive provinces.

4.3.3. Spatiotemporal Evolution of Sentiment Drivers

The sentiment driving mechanism undergoes a structural shift as the conflict evolves. We analyze the model fit and spatial autocorrelation across the policy lifecycle in Table 4.4 and Table 4.5, which reveal how the explanatory power of the thematic clusters fluctuates during different crisis phases.

Table 4.4: Model Fit and Spatial Autocorrelation (Policy Start vs. De escalation)

Stage	**Global Moran's I**	**p value**	**OLS** R^2	**GWR** R^2	**AICc**	**BW**	**N**
Policy start	0.0010	0.348	0.1340	0.1851	96.5272	20.0	29
De escalation	-0.0486	0.478	0.2582	0.3031	91.3135	20.0	29

Table 4.5: Descriptive Statistics of Local R^2 by Stage

Stage	**N**	**Mean**	**Median**	**Min**	**Max**	**Std**
Policy start	29	0.1944	0.2003	0.1309	0.2132	0.0191
De escalation	29	0.2751	0.2739	0.2060	0.3249	0.0288

Figure 4.14 illustrates the spatial distribution of local regression coefficients for the four thematic dimensions during the overall period. The results confirm that the direction and intensity of thematic influence vary significantly across provinces.

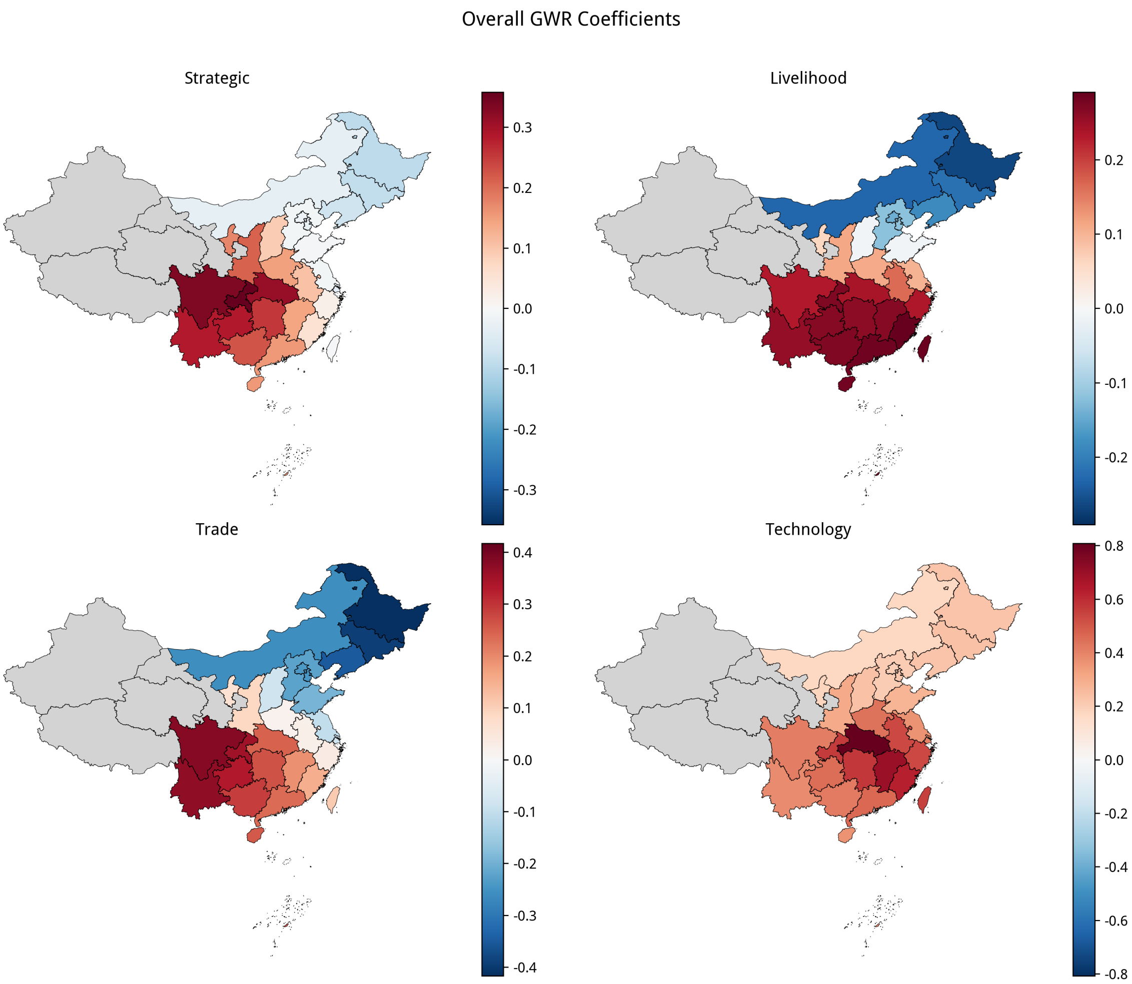

Figure 4.14: Spatial variation of local regression coefficients

Strategic topics demonstrate clear regional polarization, appearing positive in the central and western provinces but negative in the northeast. Livelihood topics exhibit a distinct north south divide, with negative influence in the north and positive sentiment correlations in the south. Similarly, trade topics appear negative in the north but transition to positive in southern clusters. In contrast, technology topics follow a positive dominant spatial pattern across the majority of provinces.

Ultimately, Figure 4.15 presents the spatiotemporal distribution of coefficients across the three policy phases.

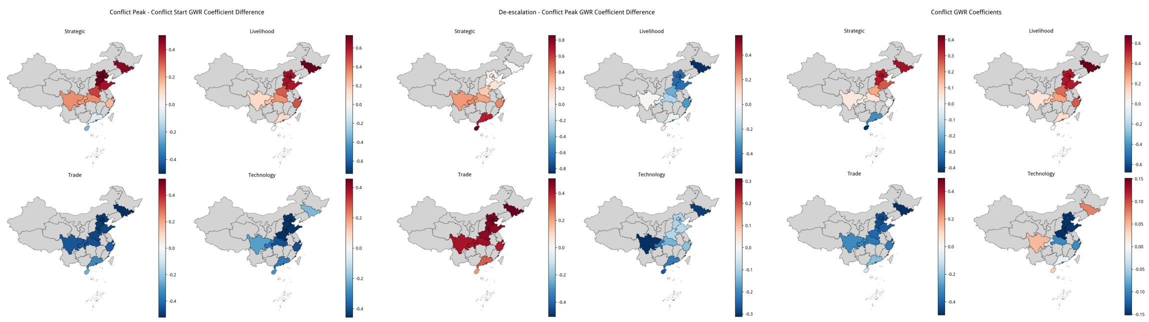

Figure 4.15: Spatiotemporal distribution of GWR coefficients across the policy lifecycle

These maps demonstrate that during the policy start phase, strategic and livelihood topics were predominantly negative, while trade and technology topics were generally positive. Upon entering the conflict escalation phase, livelihood topics transitioned to significantly positive values in most regions, while trade topics shifted toward a global negative correlation. This structural switching across the policy lifecycle highlights the profound level of spatial heterogeneity inherent in digital public responses to macroeconomic shocks.

4.4. Retweet Network Results

This section presents the empirical findings of the information cascade analysis, structured around four analytical steps: (1) establishing the structural typology of rational versus emotional cascades; (2) quantifying the role of user strata in driving cascade breadth and depth; (3) examining how sentiment polarity shapes cascade depth distributions; and (4) synthesising these findings into a unified answer to Final Research Question.

4.4.1. Structural Typology: Rational Broadcast versus Emotional Entanglement

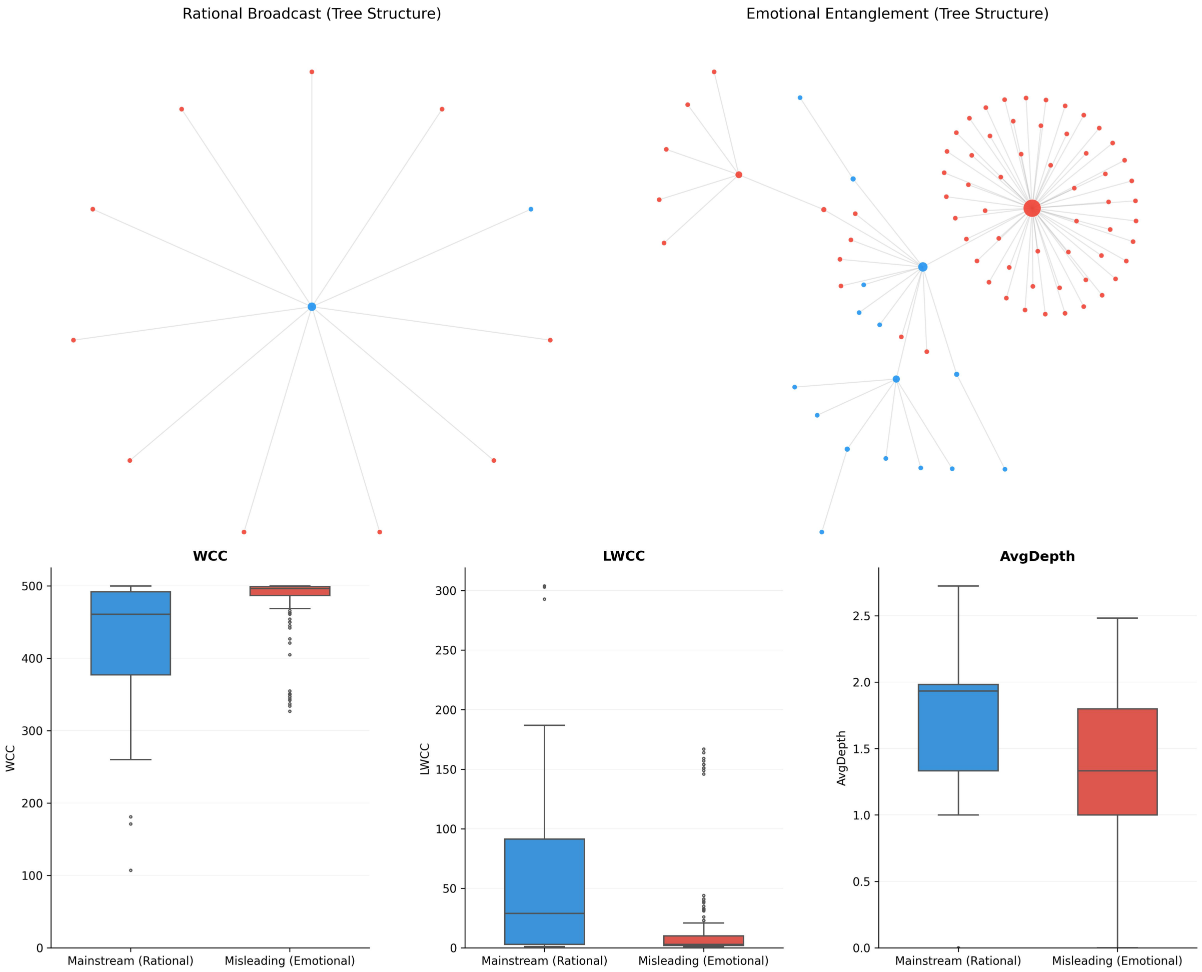

Figure 4.16: Rational vs Emotional Cascade Network Comparison

Figure 4.16 presents the topological comparison between Mainstream (Rational) and Emotionally Charged cascade networks. The tree visualisations reveal two structurally distinct diffusion archetypes. Rational cascades — originating from institutional Blue V and Red V accounts — exhibit a sparse, star-shaped broadcast topology: a single highly connected hub links directly to numerous peripheral nodes with minimal intermediate relay. Emotionally Charged cascades, by contrast, display a denser branching structure with multiple intermediate nodes absorbing and re-transmitting content across several hierarchical layers, producing what we term 'emotional entanglement.' This structural divergence is quantitatively confirmed by the three box-plot panels. On Weakly Connected Component size (WCC, an aggregate breadth measure), Rational cascades achieve substantially higher median values, indicating that institutional content commands far greater aggregate reach. The Largest Weakly Connected Component (LWCC) panel shows that Emotional cascades exhibit a wider interquartile range and more pronounced outliers: while typically smaller, emotionally charged cascades can occasionally achieve exceptional localised clustering. The Average Depth (AvgDepth) panel reveals that Rational cascades maintain a comparable or higher median depth than Emotional cascades — contradicting the intuitive assumption that emotional content propagates more deeply. This finding suggests that within Weibo's moderated environment, institutional content benefits from a structural amplification effect that simultaneously achieves wide reach and sustained penetration, while emotional content produces isolated bursts of local engagement without systematic depth.

4.4.2. User Stratum Analysis: Breadth Engines and Depth Engines

Figure 4.17 presents the Top 10 most reshared users ranked by in-degree. The ranking is dominated by Blue V institutional accounts, with two state media outlets occupying the top two positions (1,018 and 999 retweets respectively), followed by a Gold V celebrity influencer (810 retweets). The distribution then drops sharply to Yellow V commercial accounts and private users in the 250–473 range, establishing a two-tier influence hierarchy in which institutional and celebrity accounts occupy a distinct upper echelon of network centrality while commercial and ordinary accounts form a fragmented lower tier.

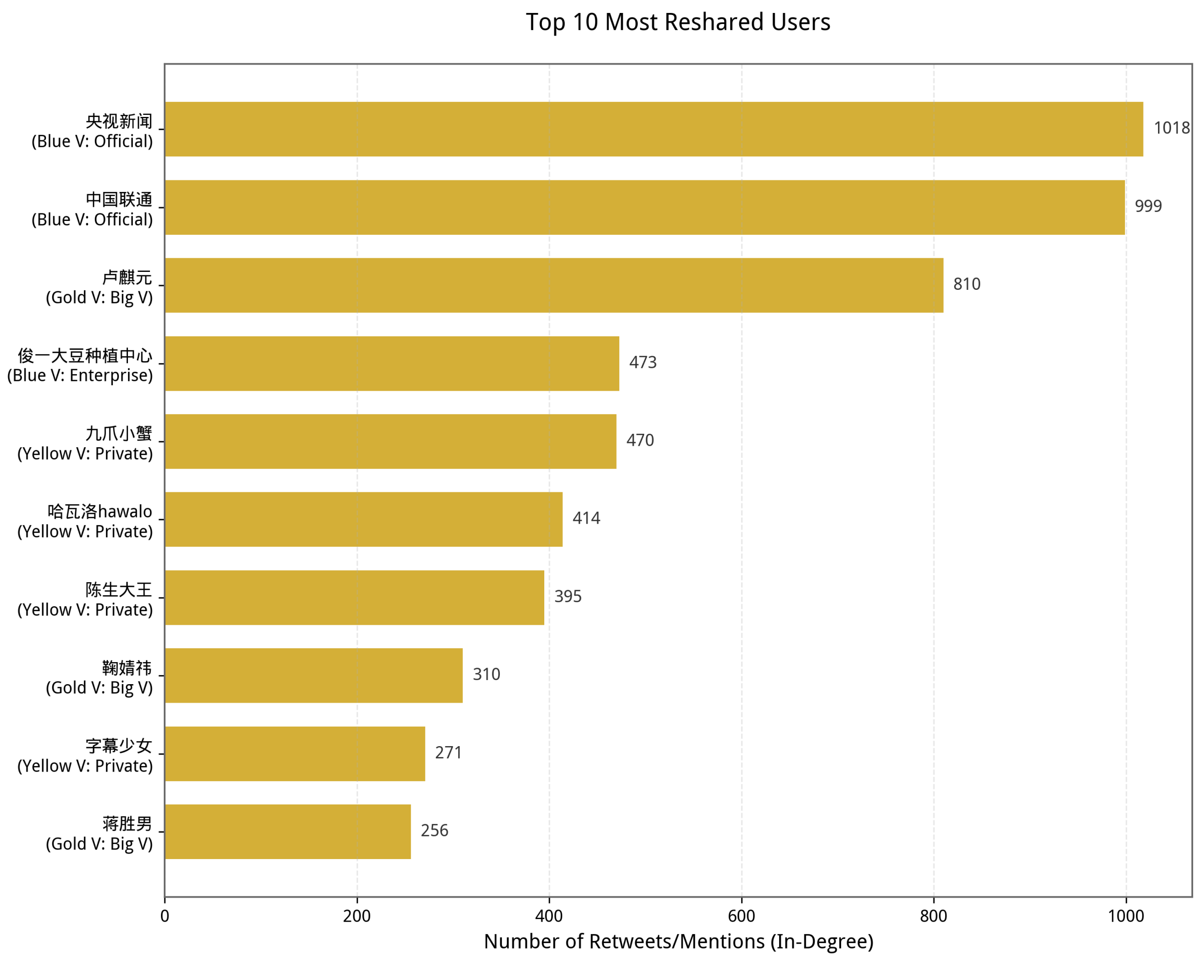

Figure 4.17: Top 10 Most Reshared Users

Table 4.6 quantifies cascade performance across all five user strata at both the Basic (Size > 2) and Impact (Size > 5) thresholds. At the Basic level, Blue V accounts generate the highest mean cascade breadth (5.92), followed by Gold V (4.72), Yellow V (4.23), Regular (3.71), and Red V (3.02). Crucially, The median cascade size is 1.0 across the full corpus (N = 552,307), and rises only marginally to 2.0 within the retweeted subset, confirming that the power-law structure of Weibo diffusion holds regardless of account type. This power-law structure confirms that meaningful diffusion is a rare, threshold-crossing event rather than the norm.

Table 4.6: Comparison of Information Cascade Metrics by User Strata

Scope	User Category	Count	Mean Br.	Median	Max Br.	Mean De.	Max De.
Basic	Gold V (金V)	2,213	4.72	2.0	1,258	1.31	4
	Red V (红V)	994	3.02	2.0	38	1.27	3
	Yellow V (黄V)	315	4.23	2.0	76	1.30	2
	Blue V (蓝V)	1,826	5.92	2.0	1,015	1.27	2
	Regular (普通)	4,715	3.71	2.0	573	1.25	4
Impact	Gold V (金V)	247	24.23	7.0	1,258	2.01	4
	Red V (红V)	103	9.80	7.0	38	2.00	2
	Yellow V (黄V)	42	17.07	11.0	76	2.00	2
	Blue V (蓝V)	134	51.36	8.0	1,015	2.00	2
	Regular (普通)	454	17.84	9.0	573	2.01	4

Note: Basic represents Size > 2; Impact represents High-Impact Cascades (Size > 5). Br. and De. denote Breadth and Depth respectively.

The Impact threshold most clearly exposes the structural role of each stratum. Blue V's mean breadth surges from 5.92 to 51.36 (an 8.7-fold amplification) once long-tail noise is filtered, ranking first by a wide margin and revealing a *super-broadcaster* dynamic: institutional accounts possess latent amplification capacity that is activated only after content crosses a minimum engagement threshold, at which point official endorsements and algorithmic traffic support rapidly build star-shaped networks of exceptional coverage. Gold V follows at 24.23, while Red V and Yellow V plateau in the 9.8–17.1 range, reflecting a clear mobilisation ceiling in the absence of either official backing or top-tier celebrity status. Figure 4.18 visually confirms this threshold activation effect, with the Basic-to-Impact contrast most pronounced for Blue V.

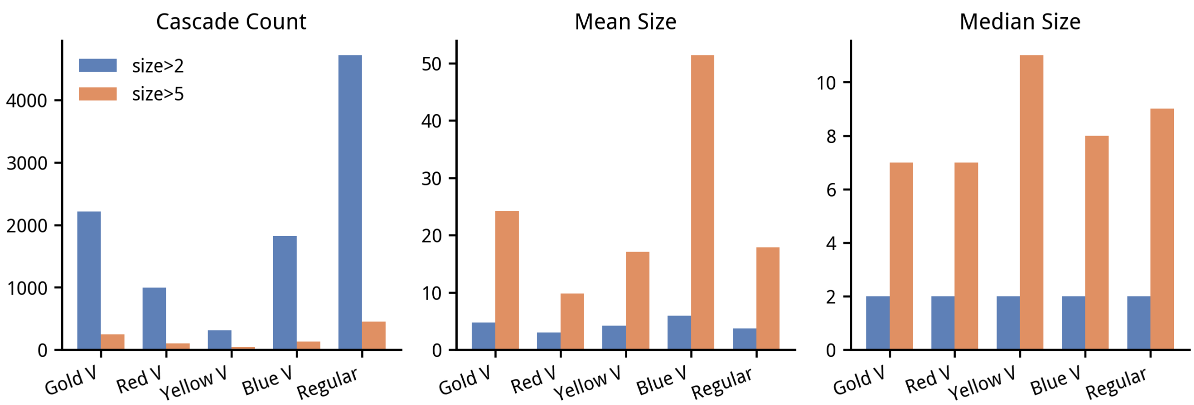

Figure 4.18: Cascade Threshold Effects by User Type (Size > 2 vs Size > 5)

The depth dimension reveals the complementary side of this architecture. Gold V influencers and Regular users achieve the highest maximum cascade depth (4 levels), while Blue V is structurally constrained to a maximum of 2. Institutional accounts therefore function as *breadth engines* — generating wide, shallow star-shaped cascades through one-to-many broadcasting — whereas elite influencers and ordinary users function as *depth engines*, driving multi-level relay cascades that sequentially penetrate audience segments unreachable by institutional broadcasts. Despite their individually weak voice (mean breadth 3.71), Regular users initiate nearly 50% of all cascade events (4,715 of 10,063), forming the indispensable connective tissue of the Weibo public opinion sphere.

4.4.3. Sentiment Polarity and Cascade Depth

Figure4.19presents the CCDF of cascade depth stratified by sentiment polarity.Across the log-log scale, the negative sentiment curve consistently maintains a higher tail probability than both neutral and positive curves at equivalent depth values. This means that for any given depth threshold x, negative-sentiment posts are statistically more likely to generate cascades exceeding x propagation layers, negative content carries a structural advantage in the depth dimension of diffusion.

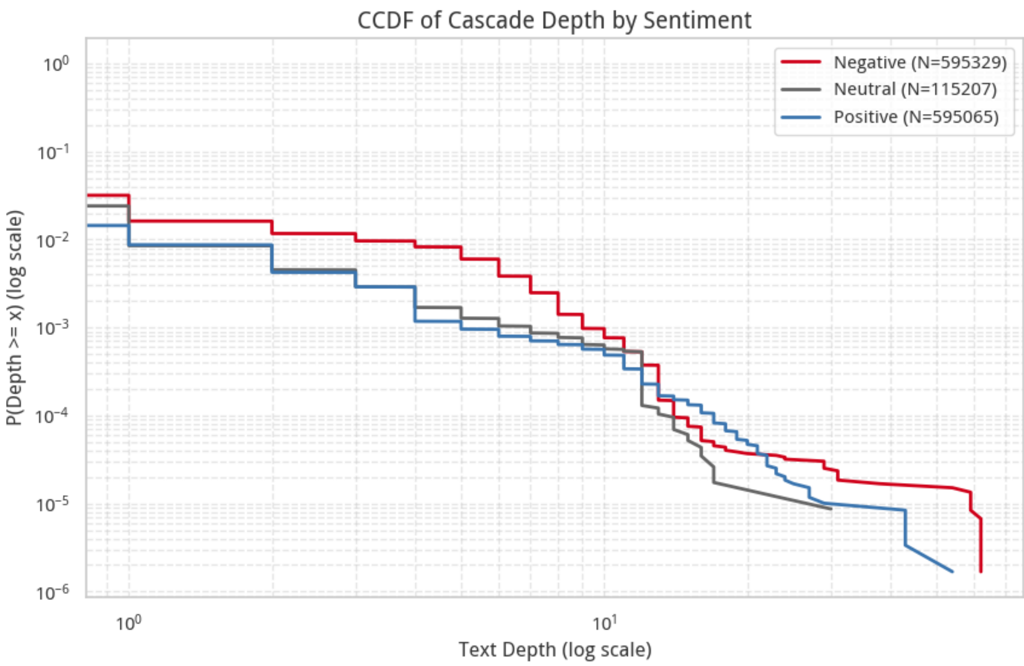

Figure 4.19: CCDF of Cascade Depth by Sentiment Polarity

CCDF results show that negative sentiment has a higher tail in the depth distribution, indicating that it is more likely to form deeper cascades. This suggests that negative content has a relative advantage in the depth dimension of propagation.

4.4.4. Synthesis: Answering RQ3

Integrating the preceding findings, user membership structures information diffusion on Weibo through role-based complementarity rather than uniform influence. Three distinct functional mechanisms emerge from the empirical evidence.

First, Blue V institutional accounts function as *breadth engines*. When their content crosses the Impact threshold, official platform endorsements and algorithmic traffic support generate star-shaped cascades of exceptional horizontal reach, making institutional broadcasting the primary mechanism of agenda-setting in tariff-related discourse. Second, Gold V celebrity influencers and Regular users function as *depth engines*. Their content — particularly when negatively valenced — generates multi-level relay cascades that penetrate audience segments unreachable by institutional broadcasts, converting broad awareness into sustained affective engagement across the network. Third, negative sentiment operates as a *structural amplifier of depth*: regardless of the initiating user type, negatively valenced posts are disproportionately likely to be relayed through multiple cascade layers, ensuring that policy-induced anxiety permeates the network more persistently than neutral or positive content.

Together, these three mechanisms explain the co-construction of Weibo's tariff-related public opinion sphere: institutional broadcasting sets the informational agenda with high breadth; grassroots and influencer relaying deepens its emotional penetration; and the inherent structural virality of negative sentiment sustains that penetration beyond the immediate event window. Digital public opinion during the 2025 tariff escalation was therefore shaped not by any single actor type, but by the structural interaction between platform hierarchy, social graph topology, and sentiment polarity.

4.5. Comparative Analysis of Spatiotemporal Sentiment and Engagement Patterns

4.5.1. Engagement Dynamics: Normalized Daily Counts

Comparing normalized daily post volumes between domestic and overseas users reveals three structural differences in how the two groups process policy events. Both cohorts exhibit highly synchronized peaks during early February, late May, and late June, confirming that major tariff announcements function as global exogenous shocks that trigger immediate cross-regional attention. Beyond these shared spikes, however, the engagement profiles diverge sharply. Overseas engagement follows a pulse-like pattern — sharper, narrower peaks with rapid post-event decay — whereas the domestic Chinese response follows a progressive trajectory, featuring a gradual pre-event build-up and a sustained post-event tail. After May 2025, this divergence deepens further: the domestic normalized baseline rises gradually, signalling a transition from event-driven discussion to sustained policy discourse, while overseas engagement remains highly volatile with frequent troughs between peaks, reflecting an audience whose attention is strictly event-contingent.

4.5.2. Emotional Landscapes: Mean Sentiment Trajectories

The sentiment trajectories reveal equally fundamental differences in emotional stability between the two groups. Domestic sentiment oscillates within a narrow band of [0.45, 0.60] throughout the study period, displaying a buffered response to policy events in which even major announcements produce smooth, gradual shifts rather than abrupt reversals. Overseas sentiment, by contrast, exhibits substantially higher amplitude and wider variance, characterised by a peak-and-trough structure in which sentiment surges and drops sharply in response to individual news cycles. This greater susceptibility to short-term narratives suggests a higher polarisation risk for overseas audiences, whereas the domestic trajectory demonstrates stronger narrative continuity and emotional equilibrium — consistent with the consolidated media environment and nationalist consensus documented in prior research on Chinese social media [**?**].

4.5.3. Sub-group Analysis: The Role of U.S. Users

Because the United States accounts for approximately one-third of the overseas sample and serves as the policy origin, a direct China–U.S. comparison is warranted. The U.S. sentiment curve exhibits the most pronounced spikes and the highest emotional tension of any sub-group examined, with amplitude substantially exceeding both the domestic trajectory and the aggregate overseas average. This confirms that U.S. users are the primary driver of overseas emotional heterogeneity. The U.S. engagement pattern is similarly distinctive, defined by short-burst dynamics — rapid escalation followed by immediate dissipation — that contrast sharply with the domestic sustained-engagement model. Together, these characteristics suggest that U.S. discourse on the tariffs is predominantly reactive and episodic, organised around specific policy announcements rather than evolving as a continuous deliberative process.

4.5.4. Synthesis of Findings for RQ2

The comparative analysis yields a consistent answer to Research Question 2. Domestic and overseas audiences respond to the same policy shocks through structurally distinct modes of engagement and affect. In terms of intensity, domestic responses are steady-state and persistent: attention accumulates gradually before major events and decays slowly afterward, sustaining a continuously elevated baseline from May onward. Overseas responses, by contrast, are event-driven and high-volatility: engagement spikes sharply at policy milestones but dissipates rapidly, leaving a low and unstable inter-event baseline. In terms of affect, domestic sentiment is characterised by narrow-band stability around a neutral-to-positive mean, while overseas sentiment — driven primarily by U.S. users — is prone to rapid swings and higher polarisation risk. Methodologically, these contrasts demonstrate that normalised engagement and mean sentiment are more informative than absolute counts for cross-regional comparison, as they isolate relative response intensity and reveal the persistence structure of public attention independent of population size differences.

5 | Conclusions and Future Work

This chapter synthesizes the core findings of the research, which investigates the multi-faceted public response to the 2025 United States tariff policies within the Weibo. By integrating high precision sentiment classification with spatiotemporal modeling and information cascade analysis, this study quantifies the interplay between macroeconomic shocks and digital societal reactions. The following sections outline the main contributions, acknowledge technical and methodological limitations, and propose strategic directions for future scholarly inquiry.

5.1. Conclusion

This study provides a multidimensional identification and formalization of public sentiment dynamics on Weibo following the 2025 US tariff policies. To achieve high precision analysis, a specialized RoBERTa model was fine tuned via knowledge distillation from DeepSeek-v3, enabling the processing of 1.3 million records with superior linguistic accuracy. Integration of temporal, spatial, and network analysis reveals that digital responses to macroeconomic shocks are structured by specific patterns of societal processing rather than stochastic fluctuations.

The temporal analysis utilizes a dynamic event detection method, where a 30 day window before and after the center date serves as a base period, and final cycle lengths are determined by a 20% peak intensity threshold. This framework demonstrates how public attention cycles expand or contract in direct response to the significance of policy updates.

Spatially, the application of Geographically Weighted Regression reveals that regional economic structures and trade dependencies act as the primary drivers of divergent emotional reactions, highlighting the critical role of geographic economic realities. Furthermore, the topological analysis exposes a structural divergence in information diffusion: institutional

Blue V accounts facilitate explosive breadth, while elite influencers and regular users drive the multilevel depth of the conversation. Ultimately, these findings prove that digital public opinion is co-constructed by geographic economic realities and social network hierarchies.

5.2. Limitations

Despite the robust performance of the analytical framework, several technical and methodological constraints must be acknowledged:

- **Data Acquisition Instability:** Anti crawling mechanisms and the random sampling nature of the Weibo crawler tools caused fluctuations in daily post volumes, potentially affecting the temporal continuity of the data stream.
- **Semantic Data Drift:** The use of snowball sampling for keyword expansion introduced instances of semantic data drift, where subsequent layers of retrieved posts occasionally deviated toward peripheral social topics.
- **Geographic Coverage and Scale:** The limited availability of geolocated IP data resulted in the exclusion of certain western provinces from the spatial analysis. Furthermore, the reliance on provincial administrative divisions introduces the risk of the Modifiable Areal Unit Problem, where intra provincial heterogeneity is smoothed over.
- **Methodological Variance:** The benefits of the spatial approach stem primarily from capturing local heterogeneity rather than global structures, and the model fitting ability varied significantly across different policy stages, showing lower explanatory power during less volatile windows.
- **Topological Constraints:** The information cascade modeling remains constrained by its reliance on visible structural metadata, failing to account for the latent influence of private sharing or offline discourse occurring outside the visible Weibo API.

5.3. Future Work

To address these limitations, several directions are proposed to enhance the reliability of spatiotemporal sentiment monitoring. The automated pipeline developed in this thesis should be deployed for long term longitudinal monitoring to track the evolution of trade conflicts from immediate shocks to permanent social states. Future models should incorporate dynamic keyword filtering using Large Language Models to maintain thematic consistency and mitigate data drift during sampling. Furthermore, the integration of external economic indicators, such as real time export volumes and regional industrial exposure, would allow for testing the predictive power of social media sentiment on actual economic performance. Finally, a comparative analysis between the 2025 tariff event and the 2018 trade war would provide valuable insights into how public sensitivity and social media network structures have evolved over time, ensuring a more robust use of digital data for monitoring global macroeconomic shifts.

Bibliography

[1] R. Al Aziz, P. Agarwal, J. McGuinness, C. L. Karmaker, and J. Zhuang. Temporal patterns and life cycle dynamics of social media user activity during disasters. *Expert Systems with Applications*, 255:124454, 2024.

[2] A. Almaslukh, A. Almaalwy, N. Allheeib, A. Alajaji, M. Almukaynizi, and Y. Alabdulkarim. Top-k sentiment analysis over spatio-temporal data. *PeerJ Computer Science*, 10:e2297, 2024. doi: 10.7717/peerj-cs.2297.

[3] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts. Everyone's an influencer: quantifying influence on Twitter. *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 65–74, 2011.

[4] F. Balfour. Social media and public opinion in China. *Contemporary Politics*, 27(3): 301–315, 2021. doi: 10.1080/13569775.2021.1917163.

[5] A. Bechini, P. Ducange, F. Marcelloni, and A. Renda. Stance analysis of Twitter users: The case of the vaccination topic in Italy. *IEEE Intelligent Systems*, 36(4): 22–30, 2021.

[6] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011. doi: 10.1016/j.jocs.2010.12.007.

[7] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. Measuring user influence in Twitter: The million follower fallacy. *ICWSM*, 10:10–17, 2010.

[8] A. S. Fotheringham, C. Brunsdon, and M. Charlton. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons, Chichester, 2003.

[9] S. Goel, A. Anderson, J. Hofman, and D. J. Watts. The structural virality of online diffusion. *Management Science*, 62(1):180–196, 2016. doi: 10.1287/mnsc.2015.2158.

[10] X. Gong, Y. Han, M. Hou, and R. Guo. Online public attention during the early days of the COVID-19 pandemic: Infoveillance study based on Baidu Index. *JMIR Public Health and Surveillance*, 6(4):e23098, 2020.

[11] D. Hanny and B. Resch. Multimodal GeoAI: An integrated spatio-temporal topic-sentiment model for the analysis of geo-social media posts for disaster management. *International Journal of Applied Earth Observation and Geoinformation*, 136:104540, 2025. doi: 10.1016/j.jag.2025.104540.

[12] J. M. Hofman et al. Integrating explanation and prediction in computational social science. *Nature*, 595:181–188, 2021.

[13] I. A. Jasim, M. K. Fileeh, M. A. Ebrahhem, L. A. Al-Maliki, S. K. Al-Mamoori, and N. Al-Ansari. Geographically weighted regression model for physical, social, and economic factors affecting the COVID-19 pandemic spreading. *Earth Systems and Environment*, 2022. doi: 10.1007/s41748-022-00323-y.

[14] G. King, J. Pan, and M. E. Roberts. How censorship in China allows government criticism but silences collective action. *American Political Science Review*, 107(2): 326–343, 2013. doi: 10.1017/S000305541300001X.

[15] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009. doi: 10.1126/science.1167742.

[16] C. Lei, A. Zhang, Q. Qi, H. Su, and J. Wang. Spatial-temporal analysis of human dynamics on urban land use patterns using social media data by gender. *ISPRS International Journal of Geo-Information*, 7(9):358, 2018.

[17] K. Lei, Y. Liu, S. Zhong, Y. Liu, K. Xu, Y. Shen, and M. Yang. Understanding user behavior in Sina Weibo online social network: A community approach. *IEEE Access*, 6:13302–13316, 2018. doi: 10.1109/ACCESS.2018.2808158.

[18] Y. Liu and S. Xu. Detecting rumors through modeling information propagation networks. *IEEE Transactions on Computational Social Systems*, 3(2):46–62, 2016.

[19] F. Pierri and S. Ceri. False news on social media: A data-driven survey. *IEEE Access*, 7:151712–151733, 2019.

[20] F. Pierri, C. Piccardi, and S. Ceri. A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter. *EPJ Data Science*, 9(1):35, 2020.

[21] F. Pierri, C. Piccardi, and S. Ceri. Topology comparison of Twitter diffusion networks effectively reveals misleading information. *Scientific Reports*, 10(1):1–10, 2020. doi: 10.1038/s41598-020-58166-5.

[22] A. Sarker, L. M. Lix, et al. Finding social media signals for monitoring the impact of COVID-19. *Journal of Biomedical Informatics*, 107:103455, 2020.

[23] W.-z. Shi, F. Zeng, A. Zhang, C. Tong, X. Shen, Z. Liu, and Z. Shi. Online public opinion during the first epidemic wave of COVID-19 in China based on Weibo data. *Humanities and Social Sciences Communications*, 9(1):1–11, 2022. doi: 10.1057/s41599-022-01181-w.

[24] X. Wang, F. Chao, and G. Yu. Evaluating rumor debunking effectiveness during the COVID-19 pandemic crisis. *Frontiers in Public Health*, 9:770111, 2021.

[25] X. Wang, B. Xia, Q. Chen, M. Skitmore, and H. Liu. Exploring public attitudes of continuing care retirement communities in China: a sentiment analysis of China's social media Weibo. *Frontiers in Public Health*, 12:1454287, 2025. doi: 10.3389/fpubh.2024.1454287.

[26] J. Yan, J. Yue, J. Zhang, and P. Qin. Research on spatio-temporal characteristics of tourists' landscape perception. *IJERPH*, 20(5):3843, 2023.

[27] H. Zhang and G. McKenzie. A spatial-temporal public opinion analysis of IP location disclosure on Chinese social media platforms. In *Proceedings of the 5th International Workshop on Location and Social Media*, 2022.

[28] B. Zhu, X. Zheng, H. Liu, J. Li, and P. Wang. Analysis of spatiotemporal characteristics of big data on social media sentiment with COVID-19 epidemic topics. *Chaos, Solitons & Fractals*, 140:110123, 2020. doi: 10.1016/j.chaos.2020.110123.

A | Appendix A

A complete list of keywords with English translations and descriptive statistics can be found in Appendix A.1

Table A.1: Keyword Distribution and Descriptive Statistics (Ranked by Post Volume)

Keyword	English Translation	Posts	Unique Users	Median Length
关税	Tariffs	131,174	28,331	147.0
股市波动	Stock Market Volatility	57,414	10,542	236.0
国际合作	International Cooperation	52,946	13,793	830.0
投资者信心	Investor Confidence	45,304	6,561	458.0
芯片	Chip	41,481	15,471	204.0
特朗普	Donald Trump	38,434	15,491	102.0
政治舆论	Political Opinion	29,719	6,992	513.0
政策反馈	Policy Feedback	29,657	6,391	928.0
跨境电商	Cross-border E-commerce	27,541	9,670	206.0
大豆	Soybeans	25,931	14,875	61.0
半导体	Semiconductor	24,230	6,729	220.0
A股	A-share Market	23,752	5,406	199.0
贸易壁垒	Trade Barriers	18,348	5,010	422.0
钢铁	Steel	18,244	10,506	79.0
民意	Public Sentiment	15,140	8,005	187.0
汽车	Automobile	12,207	8,500	83.0
协议	Agreement	10,416	5,726	284.0
电动车	Electric Vehicle	8,141	4,357	108.0
美国	USA	7,319	4,343	129.0
美元	US Dollar	5,485	2,952	185.0

Continued on next page

Table A.1 – Continued from previous page

Keyword	English Translation	Posts	Unique Users	Median Length
出口	Export	5,438	3,857	147.0
合作	Cooperation	5,415	4,455	104.0
全球经济	Global Economy	5,181	2,982	297.0
经济	Economy	4,641	3,305	123.0
贸易战	Trade War	4,198	2,465	217.0
壁垒	Barriers	4,163	2,705	133.0
贸易逆差	Trade Deficit	4,142	2,426	338.0
稀土	Rare Earth	3,729	2,372	124.0
国际资本流动	International Capital Flows	3,515	1,734	939.0
光伏	Photovoltaic	3,491	1,754	235.0
供应链	Supply Chain	3,391	1,998	329.0
关税政策	Tariff Policy	2,750	1,507	295.0
黄金	Gold	2,405	1,790	76.0
拜登	Joe Biden	2,377	1,257	195.0
欧盟	European Union	1,940	820	131.0
产业转移	Industrial Transfer	1,931	1,363	293.0
反制	Countermeasures	1,611	914	415.0
抗议	Protest	1,317	967	90.0
保护主义	Protectionism	1,151	748	511.0
301调查	301 Investigation	1,179	823	475.0
全球股市	Global Stock Markets	1,041	611	482.0
汇率	Exchange Rate	1,014	386	171.0
进口	Import	840	522	146.0
中国制造	Made in China	828	511	103.0
通货膨胀	Inflation	728	397	148.0
贸易政策	Trade Policy	661	527	370.0
外贸	Foreign Trade	628	529	109.5
进口替代	Import Substitution	445	328	755.0
双边关系	Bilateral Relations	401	130	249.0
中美达成协议	China-US Agreement	376	316	326.5
新能源	New Energy	361	334	81.0
GDP	Gross Domestic Product	358	314	168.5
全球贸易战	Global Trade War	337	116	371.0
豁免	Exemptions	337	274	223.0
反制措施	Retaliatory Measures	335	131	1,451.0

Continued on next page

Table A.1 – Continued from previous page

Keyword	**English Translation**	**Posts**	**Unique Users**	**Median Length**
多边贸易	Multilateral Trade	334	201	718.5
国产替代	Domestic Substitution	333	177	584.0
反倾销	Anti-dumping	222	157	180.0
对华关税	China Tariffs	206	121	688.5
中美贸易	China-US Trade	176	103	233.0
WTO	WTO	164	79	452.0
RCEP	RCEP	152	113	545.0
经济制裁	Economic Sanctions	111	30	599.0
汇率波动	Exchange Rate Volatility	99	64	517.0
资本外流	Capital Outflow	70	51	339.0
自由贸易区	Free Trade Zone	47	9	2,152.0

Table A.2: Descriptive Statistics for the Full Cascade Corpus Prior to Filtering

Metric	Count	Mean	Std	Min	P25	Median	P75	Max
Cascade Size (Rows)	552,307	2.36	5.15	1.0	1.0	1.0	2.0	690.0
Max Structural Depth	552,307	0.06	0.24	0.0	0.0	0.0	0.0	1.0
Max Text Depth	552,307	0.14	0.90	0.0	0.0	0.0	0.0	62.0
Max Total Implied Depth	552,307	0.14	0.90	0.0	0.0	0.0	0.0	62.0
Forward Summation	552,307	113.13	11,531.99	0.0	0.0	0.0	0.0	6,276,895.0
Comment Summation	552,307	110.78	8,548.55	0.0	0.0	0.0	1.0	3,887,476.0
Like Summation	552,307	7,304.04	131,958.40	0.0	0.0	1.0	19.0	19,999,400.0

B | Appendix B

It may be necessary to include another appendix to better organize the presentation of supplementary material.

List of Figures

List of Tables

List of Symbols

Symbol	Description	Unit / Range
Sentiment Analysis		
$\mathcal{D}_{raw}$	Full raw Weibo post corpus	—
$\mathcal{D}_{sub}$	Stratified balanced subset sampled from $\mathcal{D}_{raw}$	—
$\mathcal{D}_{silver}$	Silver-label training set generated by DeepSeek-v3	—
M_{base}	Pre-trained Chinese-RoBERTa-wwm-ext model	—
M_{tuned}	Fine-tuned sentiment classifier	—
x_i	Individual Weibo post (input text)	—
y_i	Pseudo sentiment label assigned by teacher LLM	$\{-1, 0, 1\}$
$\mathcal{L}$	Cross-entropy training loss	—
p	Predicted class probability output by M_{tuned}	$[0, 1]$
LR	AdamW optimiser learning rate	2×10^{-5}
Temporal Analysis		
V	Daily post volume sequence $\{v_1, v_2, \ldots, v_n\}$	posts/day
v_t	Post volume on day t	posts/day
V_{smooth}	7-day simple moving average of V	posts/day
V_{max}	Global peak of V_{smooth}	posts/day
τ	Peak intensity threshold ratio	0.2
H	Significance threshold $= V_{max} \times \tau$	posts/day
T_{start}	Start day of an active event phase	date
T_{peak}	Central day (local maximum) of an event phase	date
T_{end}	End day of an active event phase	date
Spatial Analysis — Autocorrelation		
G_i^*	Getis-Ord local statistic for spatial hot/cold spots	—
w_{ij}	Spatial weight between provincial units i and j	—
x_j	Attribute value (sentiment score) at unit j	$[0, 1]$
$\bar{X}$	Global mean of attribute values	$[0, 1]$
S	Global standard deviation of attribute values	—
n	Total number of spatial units (provinces)	count
I	Global Moran's I statistic	$[-1, 1]$

Summary of Technical Hyperparameters

Category	**Hyperparameter Description**	**Setting / Value**
Sentiment (RoBERTa)	Learning Rate (AdamW)	2×10^{-5}
	Batch Size / Max Sequence Length	32 / 128
	Fine tuning Epochs	5
Temporal Analysis	Simple Moving Average (SMA) Window	7 Days
	Event Window Symmetry	$\pm$15 Days
	Significance Peak Threshold (τ)	20%
Spatial Analysis (GWR)	Kernel Function	Adaptive Bi square
	Optimization Criterion	AICc
Network Cascades	Basic Engagement Threshold (Size)	> 2
	Impact Engagement Threshold (Size)	> 5

Acknowledgements

Time flies,As my time at Politecnico di Milano comes to an end, I realize how lucky I've been to meet so many people who genuinely helped me along the way.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Francesco Pierri. His unwavering commitment to academic rigor and his expansive scholarly vision have been truly transformative for me. Throughout every stage of this thesis, his guidance has been invaluable. As a newcomer of research, I am especially grateful for his patience and meticulous approach. Despite his demanding schedule, he always made time to address my questions—large or small—with promptness and care. I am deeply appreciative of the excellent research environment and facilities he provided, and I hold the utmost admiration for his academic brilliance and his generous, encouraging spirit.

I would also like to extend my sincere thanks to my co-supervisor, Geng Liu. Since our paths crossed by a stroke of luck, he has been a steady mentor on my research journey. By sharing his own experiences and insights, he has helped me navigate the challenges of the present and prepare for the future. He taught me that the consistency required for daily research is a philosophy that applies to all aspects of life—fostering a positive and disciplined attitude. During the early stages of this work, when I felt overwhelmed and uncertain, he helped me restructure my logic, guided my literature review, and emphasized the importance of daily coding practice and skill acquisition. His encouragement has been a constant source of strength.

Beyond the academic sphere, I want to thank my girlfriend, Jianing Guo, for her continuous support, wise advice, and unwavering companionship throughout this journey. My thanks also go to my friends in Milan; their presence and encouragement during high-pressure moments made all the difference.

Finally, my deepest thanks go to my family. Their unconditional love and selfless support have been the foundation upon which I have built my achievements.