

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

Jianwei Deng

`jianwei.deng@mail.polimi.it`

Politecnico di Milano

Article

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

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11104602/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

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

Anonymous author(s)

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.

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. 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 [2].

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 [3]. 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 [4]. 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 [5, 6]. 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.

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. We integrate Geographically Weighted Regression (GWR) to link spatial sentiment variance with provincial trade-to-GDP ratios [3], and utilize information cascade modeling to evaluate the structural virality of propagation across institutional and individual accounts [7, 8].

The study 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 [2]. With over 600 million monthly active users, it constitutes a unique information ecosystem shaped by specific algorithmic mechanisms and user strata [9]. 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.

1.2 Research Questions

The central research question addressed in this study is:

How do Weibo users respond to macro-level tariff policies, and what are the underlying mechanisms governing these responses across temporal, spatial, and topological network dimensions?

This objective is partitioned into three specific sub-questions:

- **RQ1 (Temporal):** How do sequential tariff shocks shape the temporal evolution, volatility, and decay rate of public sentiment?
- **RQ2 (Spatial):** How do sentiment polarities and thematic topics vary across regions, and to what extent does spatial non-stationarity reflect regional economic dependencies?
- **RQ3 (Network Topology):** How do verified user categories (institutional vs. grassroots) govern the structural reach, depth, and virality of sentiment cascades in retweet networks?

1.3 Objectives

To address the research questions systematically, this study pursues three primary objectives:

1. **Quantify Temporal Sentiment Shifts:** Measure fluctuations in emotional polarity and intensity around dynamic policy windows using smoothed trajectory modeling.
2. **Model Spatial Heterogeneity via GWR:** Capture local variations between economic dependencies (e.g., export orientation) and public sentiment, evaluating localized

goodness of fit against global baselines.

3. **Evaluate Network Diffusion Mechanisms:** Quantify hierarchical cascade metrics (breadth, depth, and structural virality) across five platform user tiers to distinguish broadcast vs. organic relay topologies.

1.4 Manuscript Structure

Figure 1 illustrates the overarching analytical workflow, linking data acquisition and distillation-based NLP classification to the tripartite analytical core: temporal event detection, spatial econometric modeling, and network cascade analysis.

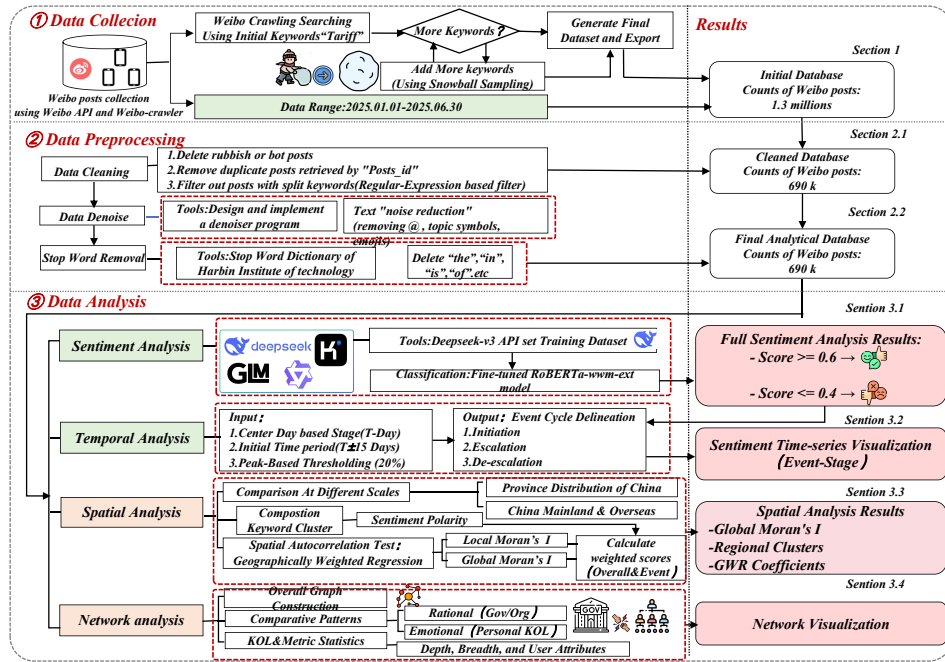


Figure 1: Integrated computational framework for multidimensional spatiotemporal and diffusion analysis.

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 and monitor public responses to external events [1]. Twitter-based studies often emphasize stance detection to separate users into opposing camps [10]. However, the discourse ecology on Sina Weibo differs significantly as national policies often show a high degree of nationalist consensus rather than structural polarization [2]. 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, providing a granular picture 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 monitoring. Temporally, collective mood follows identifiable decay patterns following external events [1, 11]. AlAziz et al. [4] demonstrated that social media reactions often exhibit multi-peak structures rather than singular isolated bursts. While previous research extensively investigated disasters or pandemic crises [11, 12], tariff policies represent deliberate, recurring geopolitical interventions. Spatially, Geographically Weighted Regression (GWR) [3] enables researchers to relax the spatial stationarity assumption. Recent studies utilized spatial analysis to map regional differences in crisis responses [13]. However, few have integrated spatial non-stationarity directly with provincial trade dependencies during international trade wars.

1.5.3 Mechanisms of Information Diffusion in Social Networks

The societal impact of an event is fundamentally moderated by its transmission topology across social hierarchies. Foundational work by Goel et al. formalized the quantification of breadth, depth, and structural virality [7]. Network topology analysis has been widely deployed to study misinformation and rumor cascades [5, 6, 8]. In contrast, the diffusion mechanics of legitimate policy sentiment across platform-certified user hierarchies remain under-explored. While follower counts do not linearly determine influence [14, 15], verified institutional accounts versus grassroots users may execute fundamentally different roles in information propagation.

1.5.4 Sentiment Analysis Methods for Social Media Text

Accurate sentiment detection in Weibo’s short, informal, and slang-heavy corpus requires advanced language representation. Pre-trained transformer architectures have significantly improved Chinese-language text classification through contextual representation and whole-word masking. Recently, spatial and multimodal sentiment analysis has advanced rapidly in social media research [16, 17]. However, direct zero-shot inference over millions of posts is computationally prohibitive. We resolve this by employing teacher-student knowledge distillation: DeepSeek-v3 generates silver annotations on a representative sample, which are subsequently used to fine-tune an efficient RoBERTa sequence classifier.

1.5.5 Summary of Research Gaps

Despite recent advancements, three critical gaps remain:

1. **Sequential Event Dynamics:** Prior work predominantly focuses on singular, sudden disaster shocks rather than dynamic phase modeling of multi-wave, intentional macroeconomic interventions.
2. **Spatial Non-Stationarity in Macroeconomics:** Most spatial sentiment studies

treat provincial sentiment as stationary or uniform, ignoring how local export dependencies induce spatial heterogeneity.

3. **Hierarchical Diffusion Roles:** The structural divergence between verified institutional broadcasters and organic grassroots relays in propagating legitimate policy emotion has not been sufficiently decomposed.

2 Data and Preprocessing

We collected over 1.3 million Weibo posts associated with 58 tariff-policy keywords between January 1 and June 30, 2025. Extracted metadata fields encompass timestamps, geographic IP tags, user verification categories, engagement metrics (reposts, comments, likes), and text strings.

2.1 Data Acquisition

2.1.1 Automated Weibo Crawler

A customized asynchronous crawler was engineered in Python to interface with the Weibo platform. The crawler gathered unique post identifiers, author UIDs, timestamps, content bodies, and IP regions. To maintain stability and prevent rate-limiting, the scraper integrated dynamic proxy rotation, exponential backoff, and header randomization.

2.1.2 Iterative Snowball Sampling

To prevent keyword selection bias and ensure the corpus expanded dynamically with the shifting public discourse, an iterative three-phase snowball strategy was implemented:

- **Phase I (Seed Identification):** Initiated with the primary seed keyword “Tariff” (关税) to establish core baseline discussions.

- **Phase II (Systematic Co-occurrence Expansion):** Top co-occurring terms and hashtags from each batch were analyzed via term-frequency analysis and filtered by domain relevance.
- **Phase III (Final Curation & Validation):** After three iterative cycles, 58 domain-specific keywords were established, covering macro-policy, affected industries (semiconductors, agriculture, automotive), and financial reactions.

Table 1 illustrates the structural schema and sample records, and Table 2 summarizes corpus dimensions.

Table 1: Data Schema and Sample Records of the Collected Dataset

Field (ID)	Description (EN)	Description (ZH)	Sample Value
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 timestamp	发布时间	2025-06-30 19:12
location	Post location	发布位置	NA
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 query tag	话题来源标签	Section 301 investigation
ip	IP region tag	IP归属地	Guangdong (广东)
retweet_id	Retweeted root ID	被转发微博ID	NA
verification	Verification status	认证类型	Gold V (金V)
membership	Membership status	会员类型	Super member
text_raw	Raw post text	原始文本	#特朗普承认美国对中国做过很多事#...

Table 2: High-level Overview of the Curated Weibo Dataset

Attribute	Value
Collection period	January 1 – June 30, 2025
Total domain keywords	58
Total original posts	1,312,847
Total retweet records	40,213
Geolocated posts (IP-tagged)	230,411
Domestic (China) geolocated posts	222,864
Overseas geolocated posts	7,547

2.2 Data Preprocessing

The text cleaning and filtering pipeline comprises four sequential stages:

1. **Spam and Bot Filtration:** Posts with empty text bodies, abnormal character-to-URL ratios, and overt advertising terms were discarded.
2. **Deduplication:** Identical posts matching across overlapping keyword queries were deduplicated based on unique post ID and SHA-256 textual hashing.
3. **Tokenization & Cleansing:** Standardized as `text_clean`, text underwent the removal of system strings, username tags (`@user`), URL entities, and graphical emoticons, converting hashtag markers (`#topic#`) to clean plain-text tokens.
4. **Corpus Diagnostics:** Figure 2 presents the volume of posts and distinct users across top keywords, verifying that data coverage was balanced without single-term monopoly.

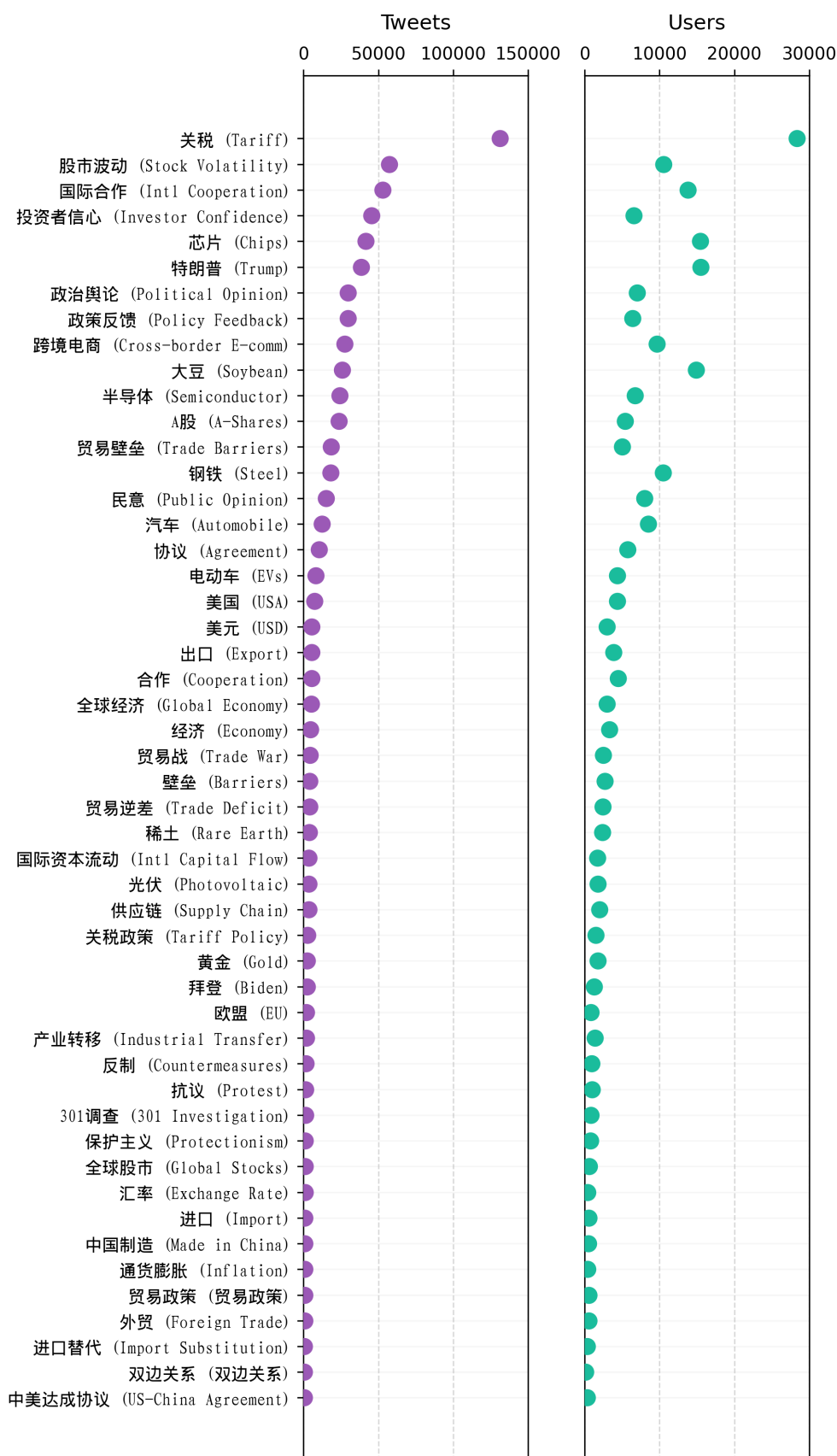


Figure 2: Number of posts (purple, left) and users (green, right) captured by each keyword in the final list (ranked by popularity) from January to June, 2025

3 Methods

This section details the methodology linking natural language processing to temporal, spatial, and network modeling.

3.1 Sentiment Classification Pipeline

To classify sentiment across 1.3 million records accurately and efficiently, we deploy a weakly supervised teacher-student framework as formalized in Algorithm 1.

Algorithm 1 Weakly Supervised Sentiment Distillation and Fine-Tuning Pipeline

Require: Raw Weibo corpus $\mathcal{D}_{raw}$, Base Transformer M_{base} , LLM Teacher (DeepSeek-v3)

Ensure: Optimized Sentiment Classifier M_{tuned}

- 1: Sample a stratified, balanced subset $\mathcal{D}_{sub} \subset \mathcal{D}_{raw}$ across keyword clusters;
 - 2: {Phase 1: Knowledge Distillation & Pseudo-Labeling}
 - 3: **for** each post $x_i \in \mathcal{D}_{sub}$ **do**
 - 4: Query teacher LLM to generate pseudo-label $y_i \in \{-1, 0, 1\}$;
 - 5: Append tuple (x_i, y_i) into silver set $\mathcal{D}_{silver}$;
 - 6: **end for**
 - 7: {Phase 2: Local Transformer Fine-Tuning}
 - 8: Initialize M_{tuned} weights from pre-trained M_{base} ;
 - 9: **while** validation loss decreases **do**
 - 10: Forward batch through M_{tuned} , computing Cross-Entropy: $\mathcal{L} = -\sum y \log(p)$;
 - 11: Backpropagate and update parameters via AdamW with learning rate LR ;
 - 12: **end while**
 - 13: **return** M_{tuned}
-

3.1.1 Teacher Selection and Model Training

We conducted an empirical validation across four LLMs (DeepSeek-v3, GLM-4.6, Qwen-Turbo, and Tongyi) on 500 representative Weibo posts. A manually verified validation split ($N = 50$) showed that DeepSeek-v3 achieved the highest classification accuracy (80.00%). DeepSeek-v3 was consequently adopted as the teacher to annotate a training set of 2,000 balanced instances.

We fine-tuned the `Chinese-RoBERTa-wwm-ext` model equipped with a classification head. Hyperparameters are reported in Table 3. The lowest validation loss (0.7551) was attained

at epoch 2 (step 226). The distilled model was subsequently used for inference across the full 1.3-million-post corpus.

Table 3: Hyperparameters for the RoBERTa Sentiment Classification Pipeline

Hyperparameter	Specification / Setting
Base Model Architecture	Chinese-RoBERTa-wwm-ext
Teacher Model	DeepSeek-v3 (API Distillation)
Optimizer	AdamW
Initial Learning Rate	2×10^{-5}
Batch Size	32
Max Token Sequence Length	128
Training Epochs	3
Loss Objective	Categorical Cross-Entropy

3.2 Temporal Dynamics and Dynamic Phase Segmentation

Rather than imposing arbitrary calendar divisions, event phases are segmented dynamically based on daily post volume variations, as detailed in Algorithm 2.

Algorithm 2 Dynamic Event Phase Segmentation

Require: Daily volume sequence $V = \{v_1, v_2, \dots, v_n\}$, threshold ratio $\tau = 0.20$

Ensure: Identified Event Phase Tuples $\{T_{start}, T_{peak}, T_{end}\}$

- 1: Compute smoothed series: $V_{smooth} = \text{SMA}(V, \text{window} = 7)$;
 - 2: Compute global maximum intensity: $V_{max} = \max(V_{smooth})$;
 - 3: Establish detection threshold: $H = V_{max} \times \tau$;
 - 4: **for** $t = 2$ to n **do**
 - 5: **if** $v_{smooth,t} > H$ **and** $v_{smooth,t-1} \leq H$ **then**
 - 6: Mark phase onset: $T_{start} = t$;
 - 7: **else if** $v_{smooth,t} < H$ **and** $v_{smooth,t-1} \geq H$ **then**
 - 8: Mark phase termination: $T_{end} = t$;
 - 9: **end if**
 - 10: **end for**
 - 11: Within each active interval $[T_{start}, T_{end}]$, locate peak $T_{peak} = \arg \max_t (V_{smooth})$;
 - 12: **return** Active intervals with corresponding symmetric analytical windows $T_{peak} \pm 15$ days;
-

This dynamic thresholding identifies three major policy lifecycle phases, summarized in Table 4.

Table 4: Dynamic Event Windows and Center Days

Phase Name	Start Date	End Date	Center Day (T_{peak})	Window ($\pm 15d$)
Policy Initiation	2025-02-06	2025-02-27	2025-02-13	Jan 29 – Feb 28
Conflict Escalation	2025-04-02	2025-04-23	2025-04-09	Mar 25 – Apr 24
De-escalation	2025-05-07	2025-05-28	2025-05-14	Apr 29 – May 29

3.3 Spatial Econometric Modeling (GWR)

3.3.1 Spatial Autocorrelation

Global Moran’s I evaluates provincial sentiment autocorrelation. To isolate localized clusters and spatial non-stationarity, Local Moran’s I_i is computed:

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

where w_{ij} represents the row-standardized spatial weight matrix between provinces i and j , x_j is the provincial sentiment mean, $\bar{X}$ is the nationwide mean, and S^2 is the global sample variance.

3.3.2 Geographically Weighted Regression Specification

GWR models localized parameters rather than assuming global stationarity:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^K \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (2)$$

where y_i is the average sentiment in province i , (u_i, v_i) denotes provincial centroid coordinates, and $\beta_k(u_i, v_i)$ represents the locally estimated regression coefficient for predictor k . The independent variables x_{ik} comprise the post volume proportions of four key thematic clusters: *Strategic*, *Trade*, *Livelihood*, and *Technology*.

An adaptive bi-square spatial kernel is calibrated by minimizing the corrected Akaike Information Criterion (AICc). Multicollinearity is checked via local Condition Numbers. Com-

putations were executed using the Python `mgwr` library.

3.4 Information Cascade and Diffusion Modeling

Retweet interactions are modeled as directed graphs $G = (V, E)$, where an edge $(u \rightarrow v)$ indicates that user u retweeted root post v . Retweet cascades are parsed into rooted propagation trees.

Users are categorized into five distinct tiers based on Weibo’s verification taxonomy:

1. **Gold V (金V)**: Verified individual influencers with high follower volumes.
2. **Red V (红V)**: Government agencies and administrative accounts.
3. **Yellow V (黄V)**: Enterprise and commercial brand entities.
4. **Blue V (蓝V)**: Official state media organizations and institutional presses.
5. **Regular (普通)**: Unverified individual citizens.

To separate trivial diffusion from high-traction propagation, two analytical filters are utilized:

- **Basic Threshold (Size > 2)**: Filters isolated dyads while preserving basic branching cascades.
- **Impact Threshold (Size > 5)**: Isolates cascades that achieve structural amplification.

Cascade breadth is defined as total nodes $|V|$, while structural depth is:

$$\text{Depth} = \max_{v \in V} d(r, v) \tag{3}$$

where $d(r, v)$ is the shortest path distance from root post r to leaf node v .

4 Results and Discussion

4.1 Macro-Level Temporal Trajectories

Figure 3 illustrates daily posting volumes alongside mean sentiment trajectories across the study period.

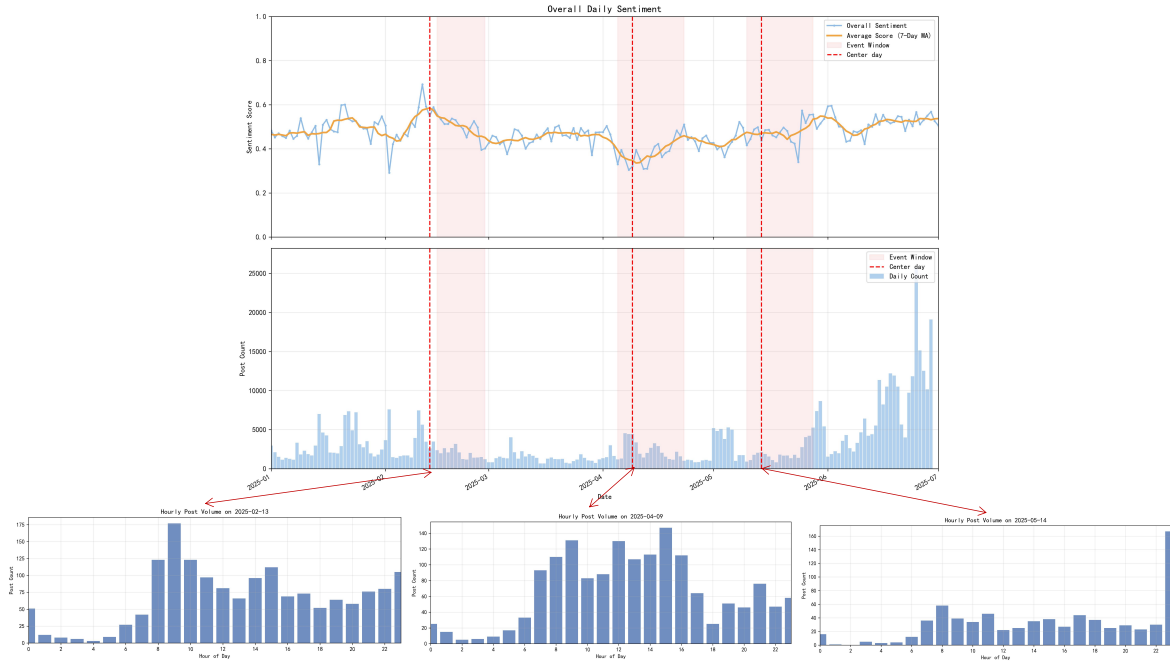


Figure 3: Temporal trends of overall sentiment and daily posting volume across event center days.

Public attention responds directly to formal policy announcements, exhibiting sharp spikes followed by gradual post-event stabilization. Figure 4 breaks down sentiment composition across the three identified policy stages.

During Phase 1 (Initiation), negative sentiment spikes rapidly due to immediate uncertainty. By Phase 2 (Conflict Escalation), overall volume reaches its peak, but discourse broadens as counter-tariffs and economic mitigation measures are announced. Phase 3 (De-escalation) demonstrates an emotional recovery toward the baseline.

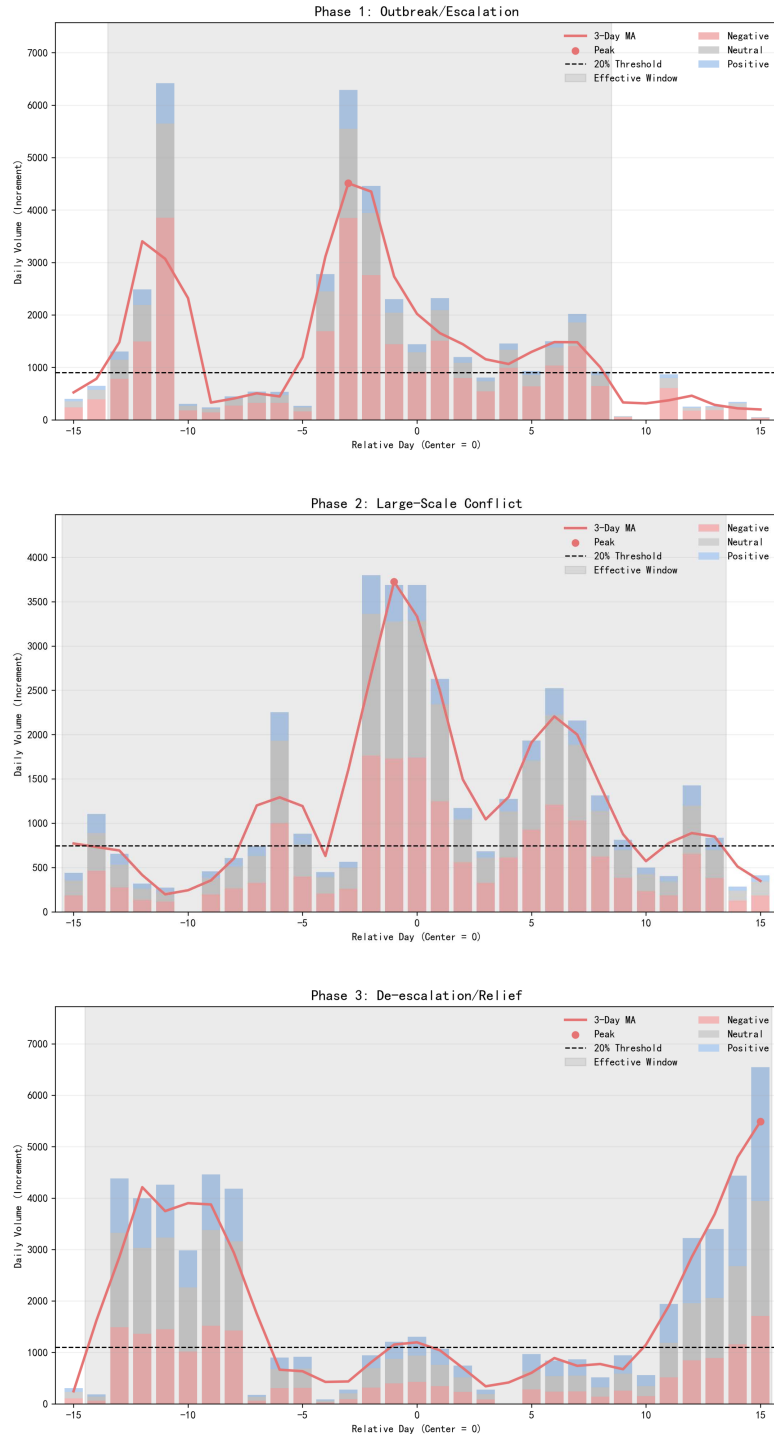


Figure 4: Sentiment polarity composition and relative volume increment across dynamic event phases.

4.2 Cross-Regional and Cross-Platform Dynamics

4.2.1 Volume and Volatility: Domestic Weibo vs. Overseas Discourse

Comparing domestic Weibo discussions to overseas discussions (Figure 5 and Figure 6) reveals distinct structural properties.

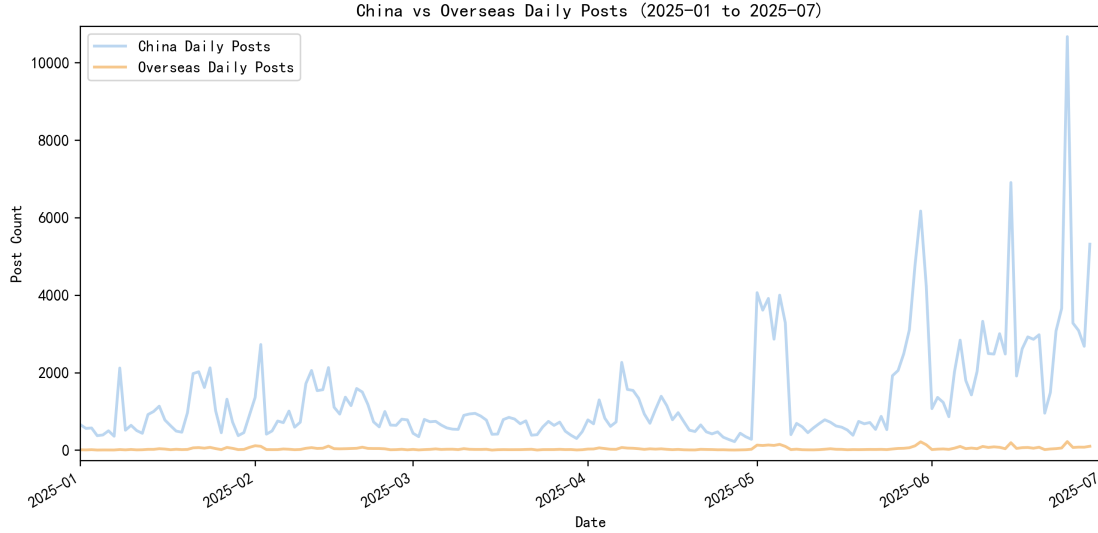


Figure 5: Absolute daily posting volume: China (Weibo) vs. Overseas.

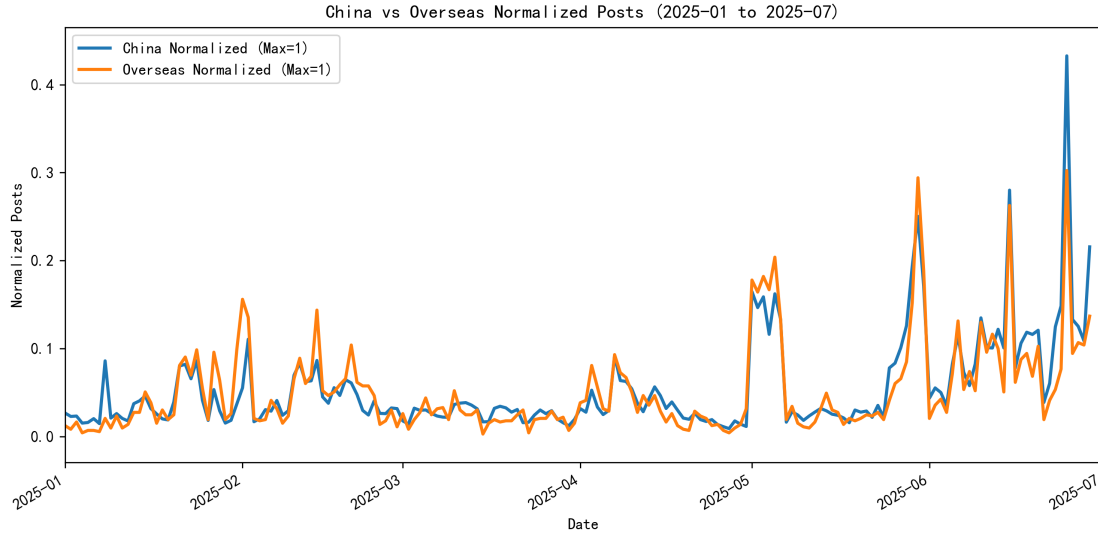


Figure 6: Normalized daily posting volume: China vs. Overseas.

While both domestic and overseas curves peak synchronously around major announcements, their response shapes diverge: overseas engagement displays an episodic *pulse-like* pattern

characterized by abrupt spikes and immediate decay. Domestic discourse exhibits a *progressive* build-up and a sustained post-event tail, establishing an elevated baseline from May 2025 onward.

4.2.2 Sentiment Trajectories: China vs. Overseas

As shown in Figure 7, domestic sentiment remains buffered within a narrow range ($[0.45, 0.60]$). Overseas sentiment exhibits substantially wider swings and heightened volatility.

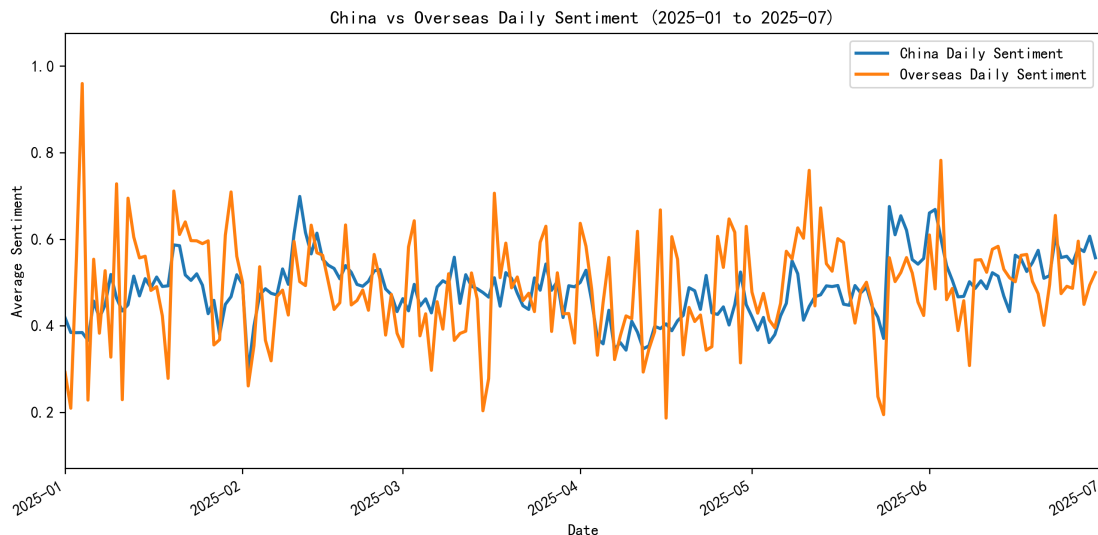


Figure 7: Daily sentiment trajectory comparison: China vs. Overseas.

4.2.3 The China–U.S. Dyad

Focusing on the U.S. subset (accounting for 30.4% of overseas data) highlights these behavioral differences (Figure 8 and Figure 9). U.S. discourse shows rapid, reactive bursts with the highest emotional amplitude among all sampled groups.

4.2.4 Geographic Distribution of Overseas and Domestic Engagement

Figure 10 and Figure 11 summarize the overseas post distribution by country and per capita. Figure 12 and Figure 13 map domestic engagement across Chinese provinces.

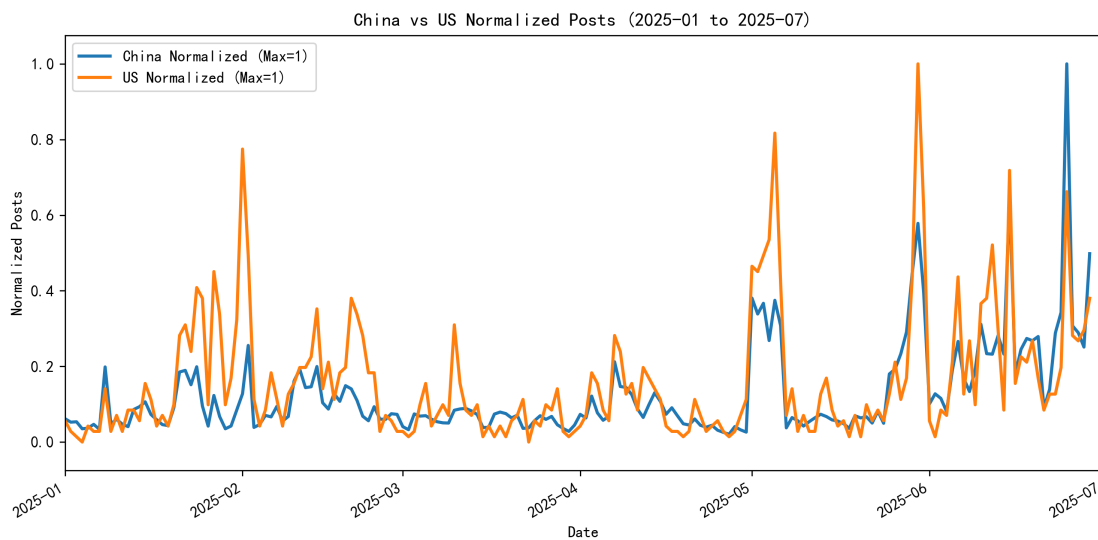


Figure 8: Normalized daily engagement volume: China vs. United States.

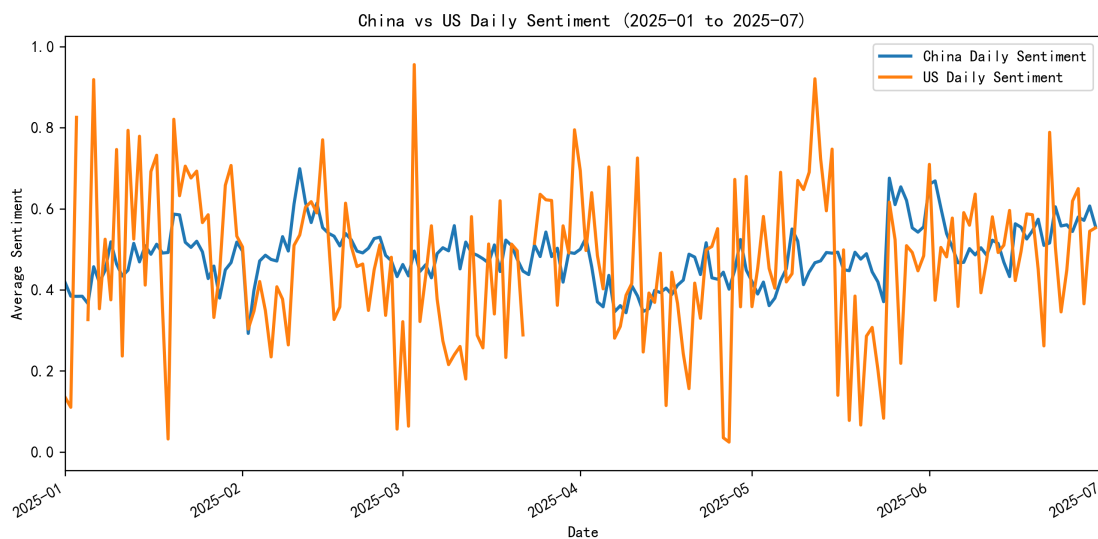


Figure 9: Daily sentiment trajectory comparison: China vs. United States.

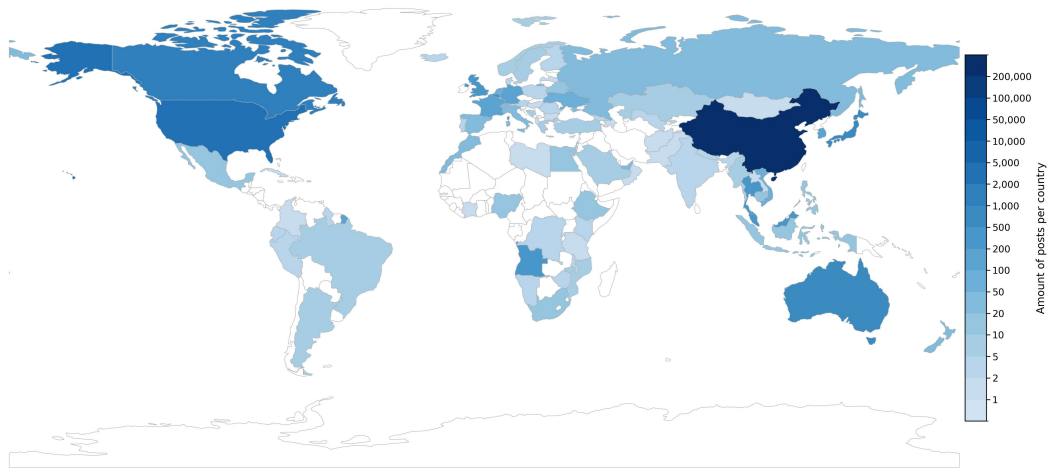


Figure 10: Total post volume by overseas country.

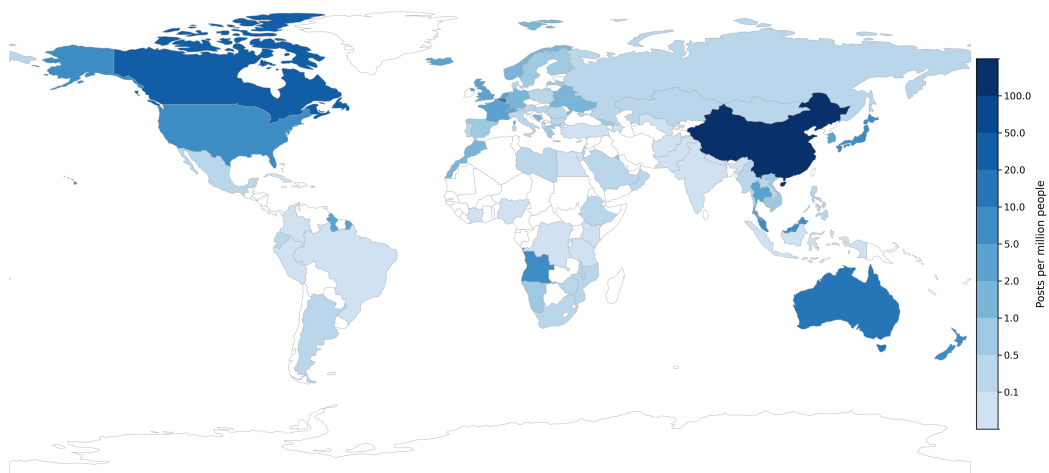


Figure 11: Posts per million residents across overseas countries.

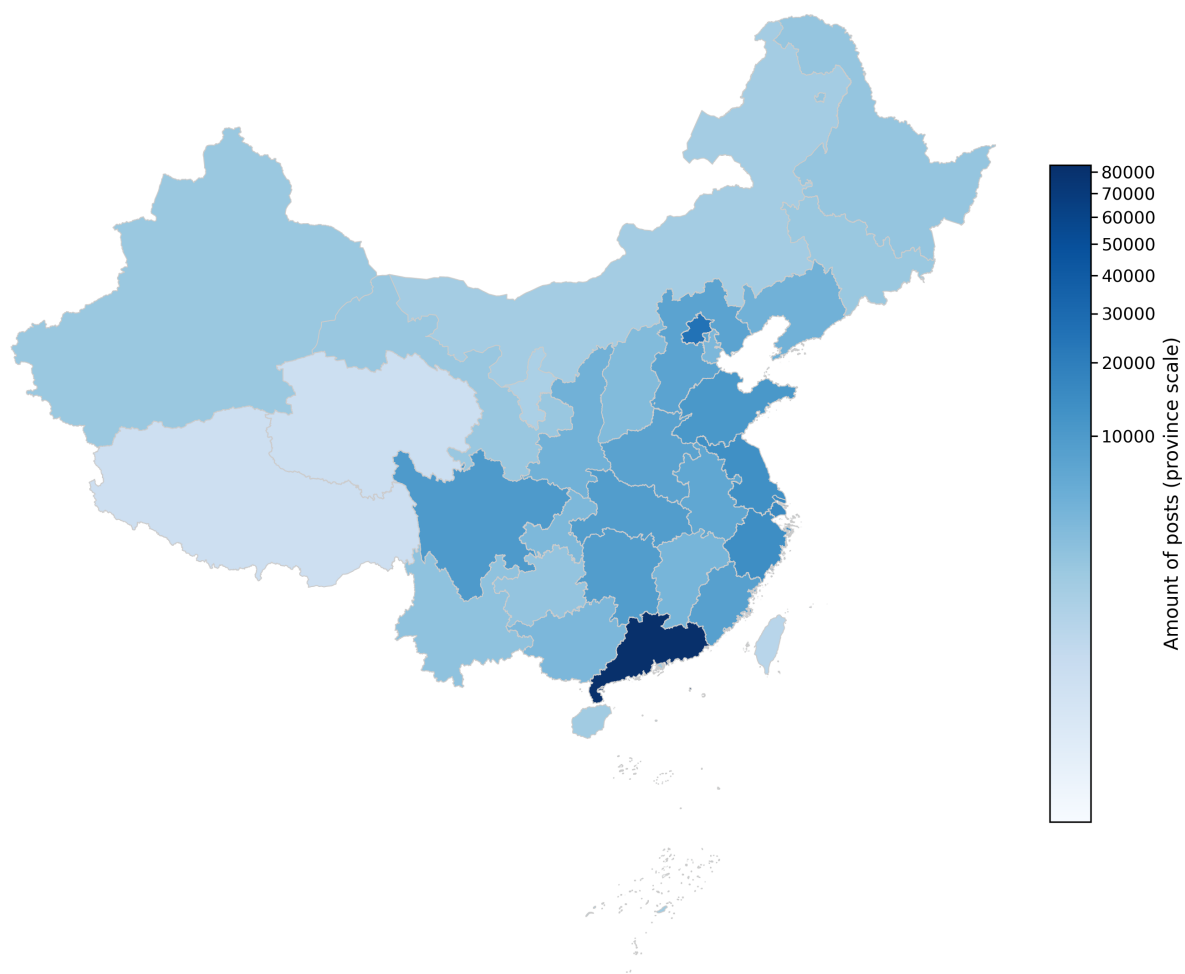


Figure 12: Spatial Distribution of Total Weibo Post Volume Across Chinese Provinces

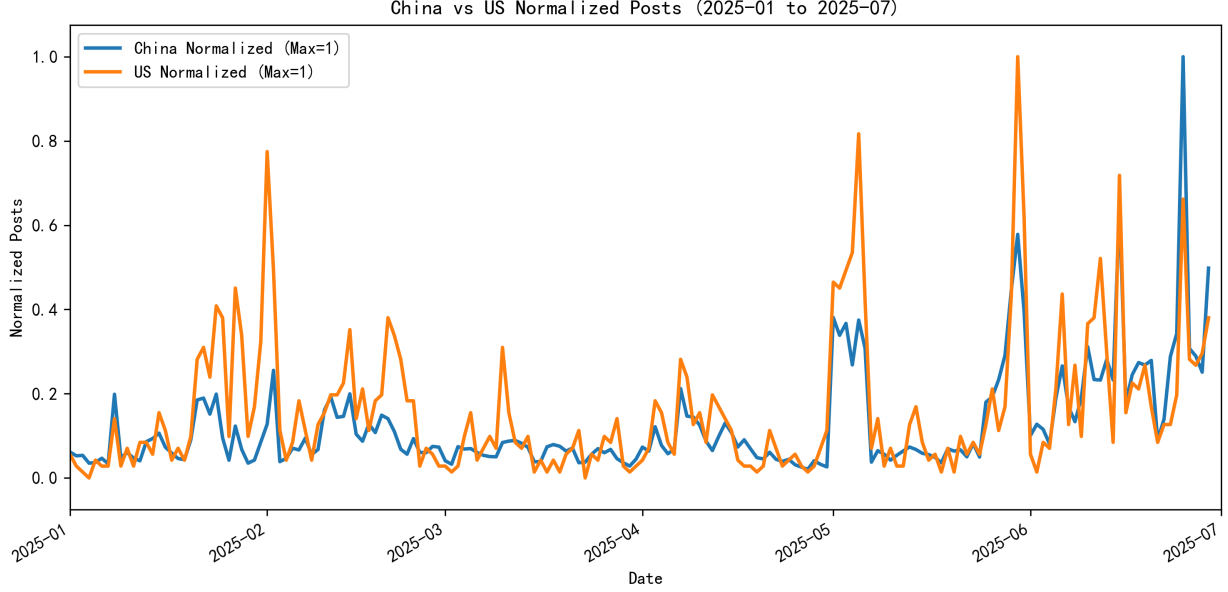


Figure 13: Comparative Temporal Trajectory of Normalized Daily Post Volume Between China and the United States

Absolute post volume is concentrated in coastal industrial hubs (Guangdong, Jiangsu, Zhejiang). Normalized per capita engagement highlights heightened sensitivity in trade-exposed coastal centers and political hubs (Beijing, Shanghai), demonstrating the spatial non-stationarity of policy salience.

4.3 Spatial Econometric Analysis (GWR Results)

4.3.1 Thematic Baseline

Table 5 summarizes post frequencies and mean sentiment scores for the four thematic clusters.

4.3.2 GWR Model Validation

Table 6 compares GWR performance with the global Ordinary Least Squares (OLS) baseline. Across the full study period, GWR increases the explained variance (R^2) from 0.0306 to 0.6004, and from 0.2570 to 0.6086 during the Conflict Escalation phase.

Table 5: Descriptive Statistics of Keyword Clusters and Sentiment Means

Cluster	Representative Keywords	Frequency	Mean Sentiment	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	
Trade	Tariffs (关税)	193,876	0.52	264,389
	Export (出口)	54,292	0.49	
	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	
Technology	Chips (芯片)	62,843	0.48	196,348
	Automobile (汽车)	60,930	0.50	
	Semiconductors (半导体)	42,730	0.58	

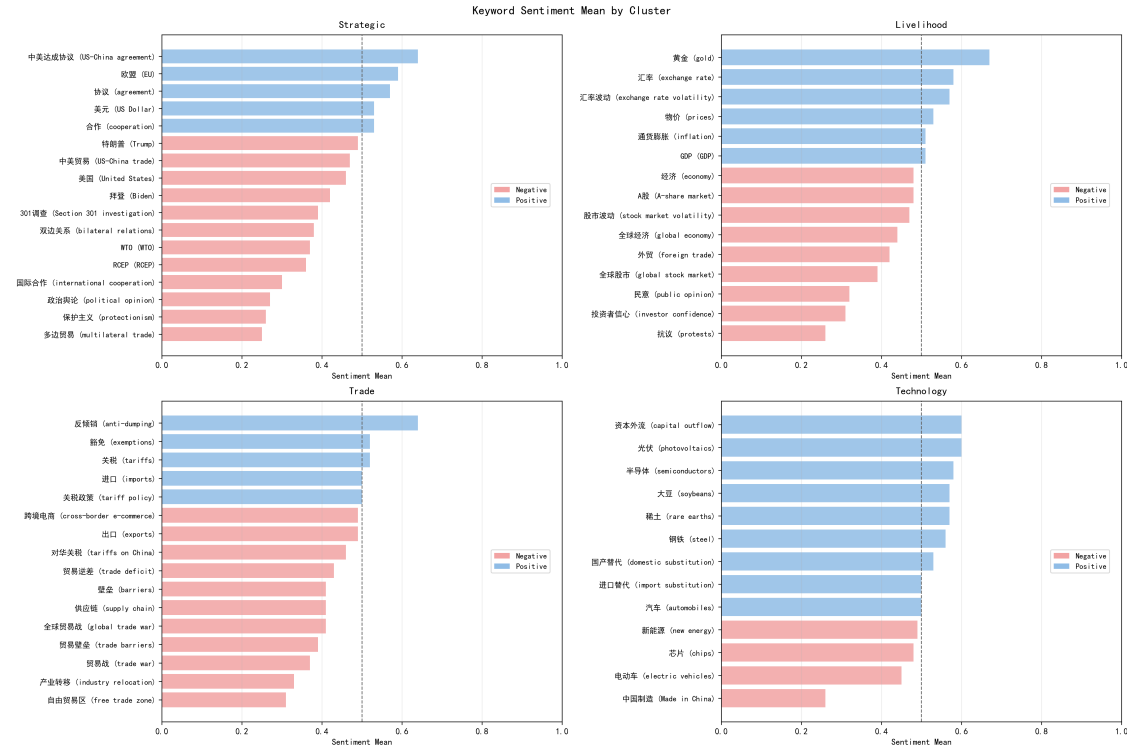


Figure 14: Comparative sentiment polarity distribution across the four thematic dimensions.

Table 6: Model Diagnostics: Global OLS vs. Local GWR

Stage	Moran's I	p -value	OLS R^2	GWR R^2	AICc	Bandwidth (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

Table 7 and Figure 15 report the spatial distribution of local R^2 values, illustrating that model explanatory power peaks across eastern coastal and northeastern provinces ($R^2 > 0.60$).

Table 7: Summary Statistics for Local R^2 (Overall Model)

Count (N)	Mean	Median	Min	Max	Std. Dev.
30	0.5601	0.5650	0.4192	0.6663	0.0731

4.3.3 Spatiotemporal Parameter Transitions

Table 8 and Table 9 track GWR performance across lifecycle phases. Spatial coefficients for each theme are plotted in Figure 16 and Figure 17.

Table 8: GWR Performance Diagnostics Across Lifecycle Phases

Stage	Moran's I	p -value	OLS R^2	GWR R^2	AICc	BW	N
Policy Initiation	0.0010	0.348	0.1340	0.1851	96.53	20.0	29
De-escalation	-0.0486	0.478	0.2582	0.3031	91.31	20.0	29

Table 9: Summary Statistics for Local R^2 by Lifecycle Stage

Stage	N	Mean	Median	Min	Max	Std. Dev.
Policy Initiation	29	0.1944	0.2003	0.1309	0.2132	0.0191
De-escalation	29	0.2751	0.2739	0.2060	0.3249	0.0288

During Policy Initiation, discussions centered on trade and strategy were negatively correlated with sentiment in export-reliant coastal hubs. In the Conflict Escalation phase, thematic coefficients shift: livelihood discussions correlate positively with sentiment as domestic economic interventions are emphasized, whereas trade discussions remain negatively correlated in northern manufacturing centers.

4.4 Retweet Network Diffusion Results

4.4.1 Topological Archetypes: Rational Broadcast vs. Emotional Entanglement

Figure 18 compares the structural topologies of Mainstream (Rational) cascades against Emotionally Charged cascades.

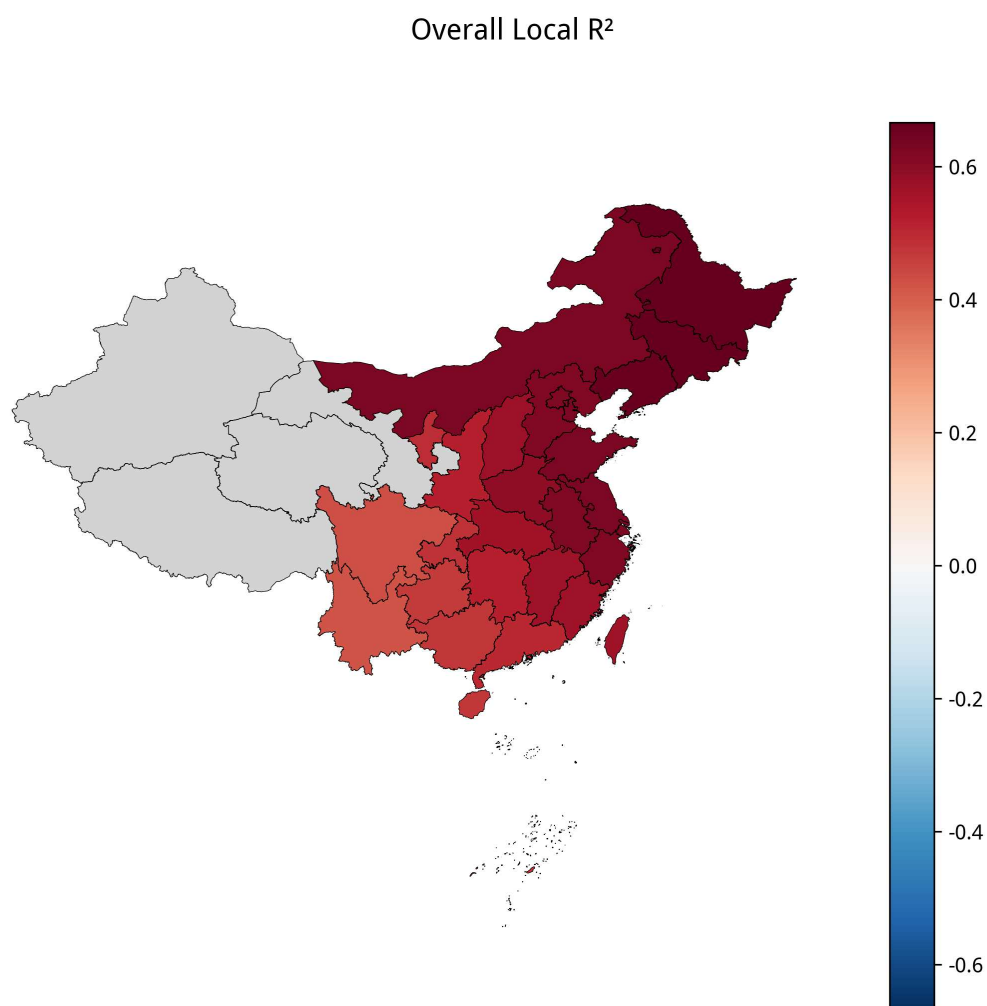


Figure 15: Spatial distribution of local R^2 goodness-of-fit values across provinces.

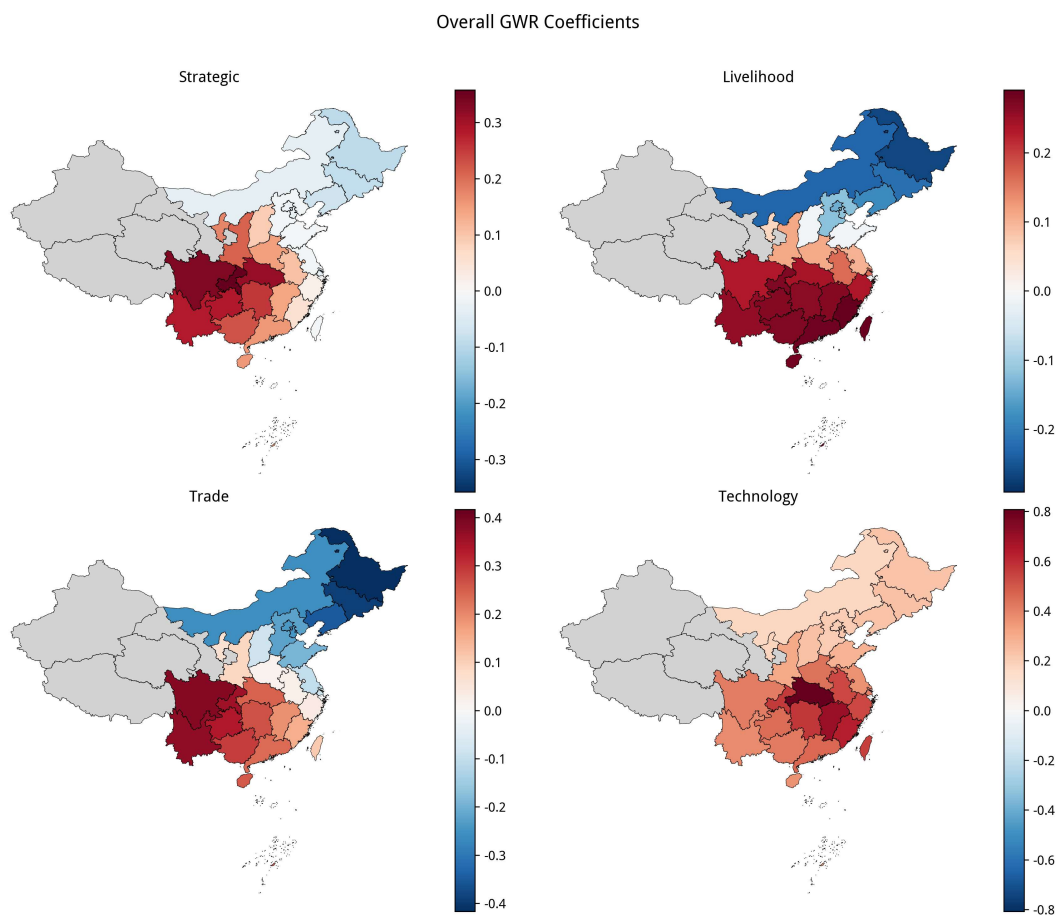


Figure 16: Spatial variation of local GWR regression coefficients across predictors.

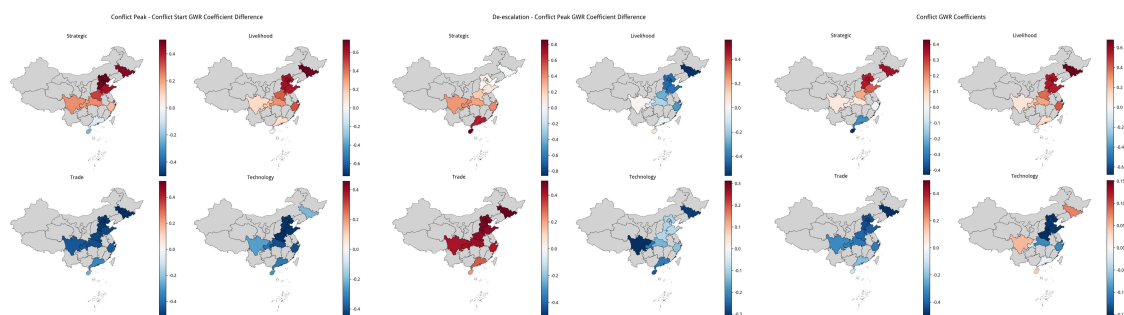


Figure 17: Spatiotemporal evolution of local GWR regression coefficients across policy stages.

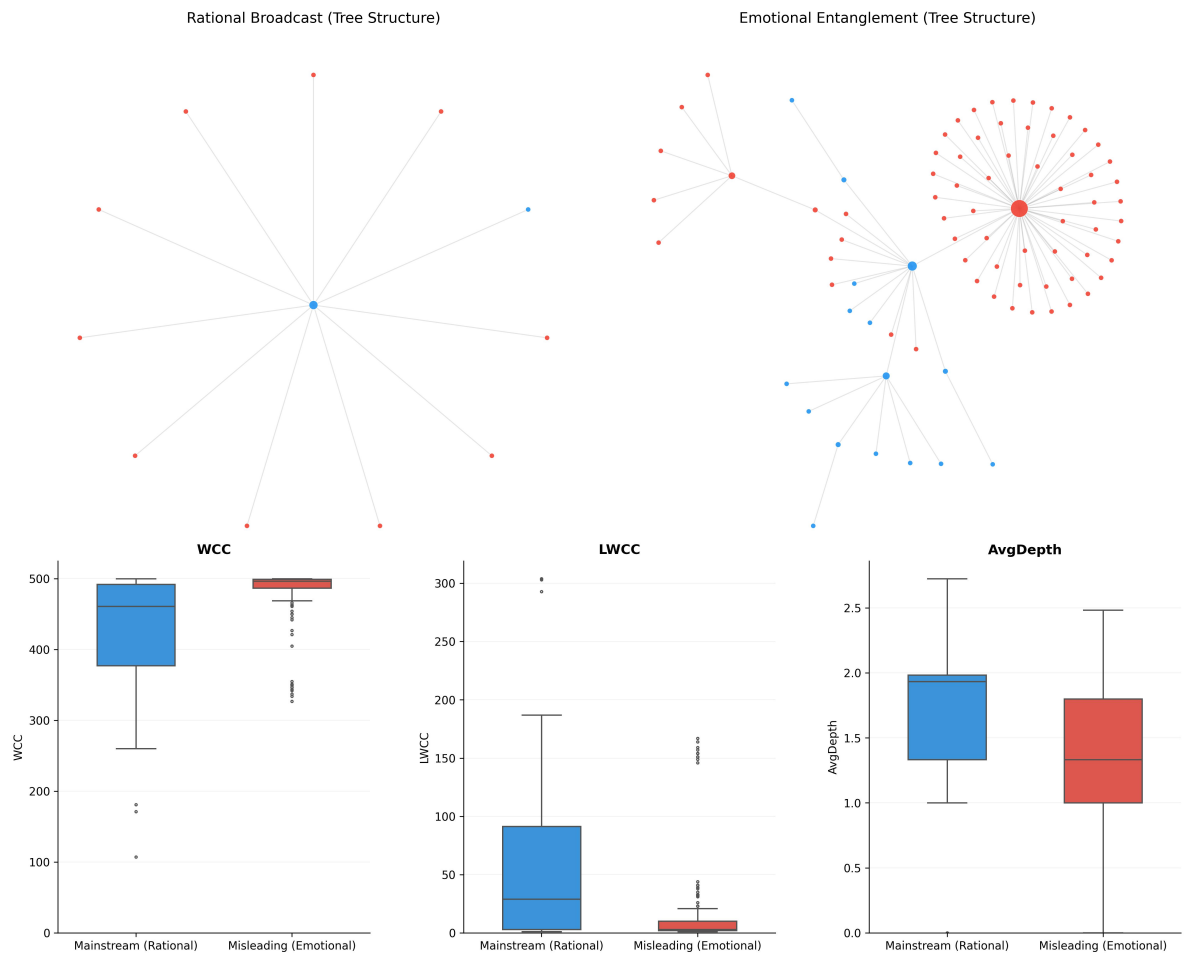


Figure 18: Topological comparison of information cascades: Rational vs. Emotional.

Rational cascades originate primarily from institutional media (Blue V) and official agencies (Red V). They exhibit a star-like topology: a central broadcasting hub connects directly to first-degree retweeters with few intermediate relays. In contrast, Emotionally Charged cascades form multi-tiered branching trees where intermediate users repost and annotate content sequentially (“emotional entanglement”).

4.4.2 User Stratum Decomposition: Breadth vs. Depth Engines

Figure 19 displays the top 10 reshared accounts. Table 10 summarizes cascade performance across user categories at both the Basic and Impact thresholds.

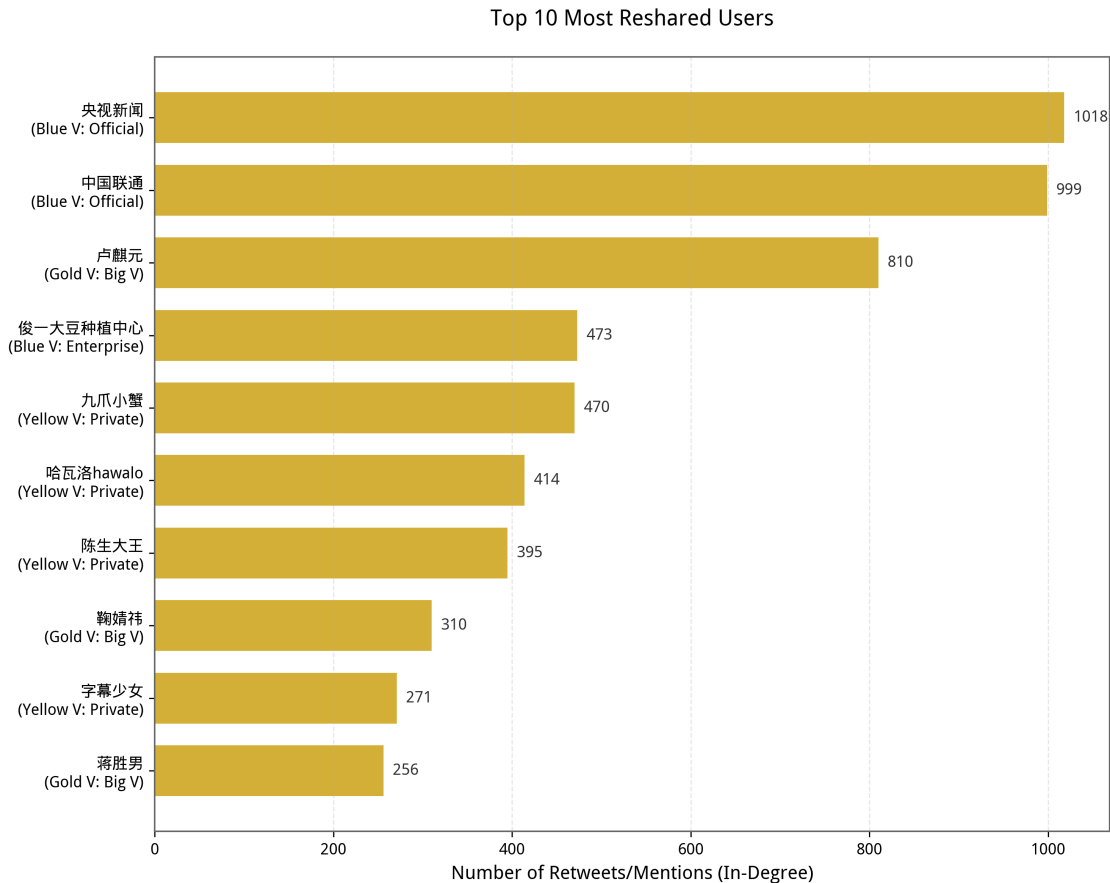


Figure 19: Top 10 most reshared accounts ranked by in-degree.

At the Impact threshold, Blue V institutional accounts exhibit an 8.7-fold increase in mean breadth (from 5.92 to 51.36), identifying them as *breadth engines*. However, their maximum

Table 10: Cascade Metrics Across User Tiers Under Basic (> 2) and Impact (> 5) Thresholds

Threshold	User Tier	Count	Mean Breadth	Median	Max Breadth	Mean Depth	Max Depth
Basic (> 2)	Gold V	2,213	4.72	2.0	1,258	1.31	4
	Red V	994	3.02	2.0	38	1.27	3
	Yellow V	315	4.23	2.0	76	1.30	2
	Blue V	1,826	5.92	2.0	1,015	1.27	2
	Regular	4,715	3.71	2.0	573	1.25	4
Impact (> 5)	Gold V	247	24.23	7.0	1,258	2.01	4
	Red V	103	9.80	7.0	38	2.00	2
	Yellow V	42	17.07	11.0	76	2.00	2
	Blue V	134	51.36	8.0	1,015	2.00	2
	Regular	454	17.84	9.0	573	2.01	4

structural depth is capped at 2 levels. Conversely, Gold V influencers and Regular unverified users attain depths of 4 levels, functioning as *depth engines*. Regular users account for 46.9% of all active cascades ($N = 4,715$), forming the structural substrate of the network. Figure 20 illustrates this threshold activation effect.

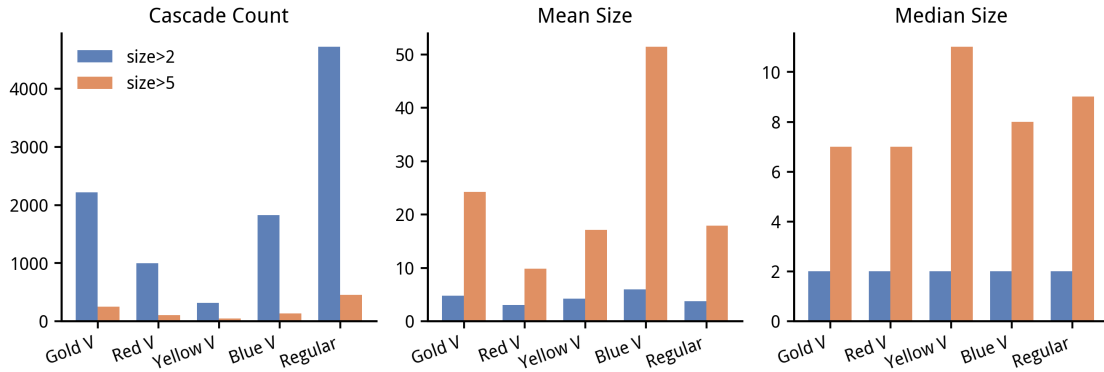


Figure 20: Cascade breadth and depth scaling by user tier across thresholds.

4.4.3 Sentiment Polarity and Diffusion Depth

Figure 21 displays the Complementary Cumulative Distribution Function (CCDF) of cascade depth stratified by sentiment polarity.

Negative sentiment exhibits a heavier tail across deeper propagation levels. For any given depth x , negative posts have a higher probability of sustaining further propagation than neutral or positive content.

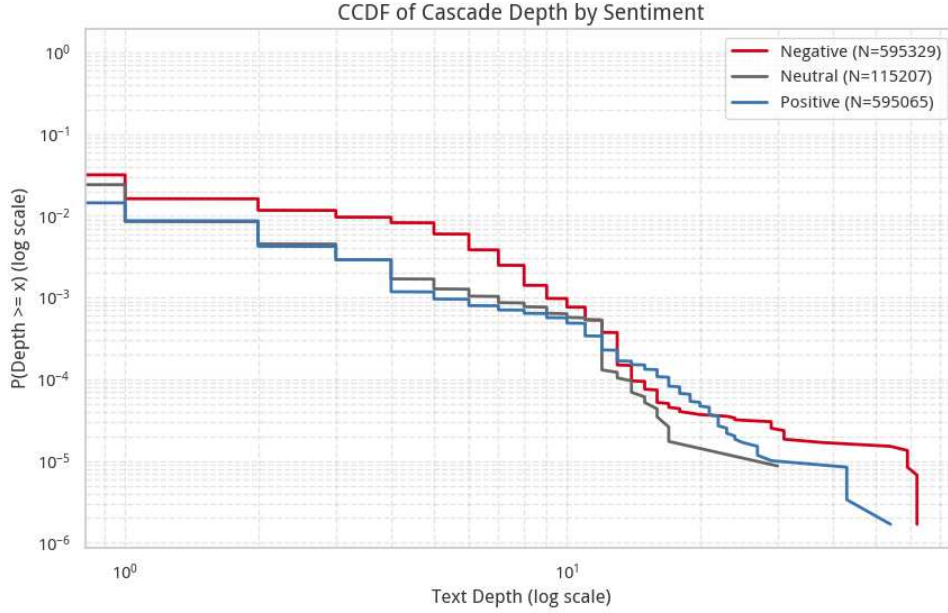


Figure 21: Complementary Cumulative Distribution Function (CCDF) of cascade depth across sentiment polarities.

4.4.4 Synthesis: Addressing RQ3

The empirical evidence indicates that information diffusion is governed by a functional division of labor:

1. **Institutional Breadth Engines:** Blue V accounts generate wide, shallow star topologies that establish agenda-setting reach.
2. **Grassroots Depth Engines:** Regular users and Gold V influencers sustain propagation through multi-level relay chains.
3. **Negative Affective Amplification:** Negative sentiment acts as a catalyst for depth penetration, sustaining propagation across intermediate nodes.

5 Conclusions and Future Work

5.1 Conclusion

This study established an integrated computational framework to examine spatiotemporal sentiment dynamics and network diffusion on Sina Weibo during the 2025 U.S. tariff impositions. By fine-tuning a `Chinese-RoBERTa-wwm-ext` classifier via knowledge distillation from DeepSeek-v3, 1.3 million posts were processed with high precision.

Temporally, public attention was delineated into three policy phases using dynamic moving-average thresholding. Cross-platform comparisons revealed that while overseas discussions followed short-lived, volatile bursts, domestic Chinese discourse exhibited buffered sentiment stability and sustained engagement. Spatially, Geographically Weighted Regression captured regional non-stationarity, demonstrating that local economic factors and export dependency explained up to 60.8% of sentiment variance. Topologically, information diffusion exhibited a clear functional split: institutional accounts served as breadth engines through star-like broadcasting, while grassroots users and individual influencers functioned as depth engines, with negative sentiment increasing structural cascade depth.

5.2 Limitations

Several limitations should be considered:

- **Crawler Sampling Constraints:** Platform rate-limiting and anti-scraping countermeasures may introduce localized sampling variances.
- **Snowball Lexicon Drift:** Iterative keyword expansion can capture tangential social discourse alongside core policy discussions.
- **Geographic Granularity:** IP resolution was constrained to provincial administrative units, smoothing over intra-provincial disparities (Modifiable Areal Unit Problem).

- **Dark Social Diffusion:** Cascade reconstruction is limited to publicly visible retweets, omitting private forwards, chat groups, and offline discussions.

5.3 Future Work

Future research directions include:

- Deploying continuous longitudinal pipelines to monitor post-shock tariff adaptations over multi-year windows.
- Integrating firm-level financial disclosures and customs datasets directly into spatial econometric models.
- Incorporating multimodal data (images, short video audio) into the classification pipeline.
- Conducting comparative cascade analyses between the 2018–2019 trade war and the 2025 escalation to evaluate shifts in platform dynamics over time.

Data Availability

Due to repository file size thresholds, the aggregated empirical datasets supporting the findings of this study—including temporal sentiment indices, provincial economic indicators for Geographically Weighted Regression (GWR), cascade network topology distributions, and variable definitions—have been deposited in a dedicated open-access storage repository: <https://drive.google.com/drive/folders/1shx9944gm3ysknShThDmxLEDFcU-Wezk?usp=sharing>

In accordance with platform terms of service and user privacy regulations, raw user-level microblog text records are withheld; however, comprehensive aggregated tabular datasets are made fully accessible for evaluation and replication.

Code Availability

All underlying computational scripts, including the RoBERTa distillation pipeline, moving-average dynamic phase segmentation, Geographically Weighted Regression (GWR), and information cascade modeling, are openly accessible in the public repository:

<https://github.com/Simon0619/Weibo-tariff-dynamics-diffusion>

References

- [1] Bollen, J., Mao, H. & Zeng, X. Twitter mood predicts the stock market. *Journal of Computational Science* **2**, 1–8 (2011).
- [2] King, G., Pan, J. & Roberts, M. E. How censorship in China allows government criticism but silences collective expression. *American Political Science Review* **107**, 326–343 (2013).
- [3] Fotheringham, A. S., Brunson, C. & Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships* (John Wiley & Sons, Chichester, 2003).
- [4] Al Aziz, R., Agarwal, P., McGuinness, J., Karmaker, C. L. & Zhuang, J. Temporal patterns and life cycle dynamics of social media user activity during disasters. *Expert Systems with Applications* **255**, 124454 (2024).
- [5] Pierri, F., Piccardi, C. & Ceri, S. Topology comparison of Twitter diffusion networks effectively reveals misleading information. *Scientific Reports* **10**, 1–10 (2020).
- [6] Pierri, F., Piccardi, C. & Ceri, S. A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter. *EPJ Data Science* **9**, 35 (2020).

- [7] Goel, S., Anderson, A., Hofman, J. & Watts, D. J. The structural virality of online diffusion. *Management Science* **62**, 180–196 (2016).
- [8] Pierri, F. & Ceri, S. False news on social media: A data-driven survey. *IEEE Access* **7**, 151712–151733 (2019).
- [9] Lei, K. *et al.* Understanding user behavior in Sina Weibo online social network: A community approach. *IEEE Access* **6**, 13302–13316 (2018).
- [10] Bechini, A., Ducange, P., Marcelloni, F. & Renda, A. Stance analysis of Twitter users: The case of the vaccination topic in Italy. *IEEE Intelligent Systems* **36**, 22–30 (2021).
- [11] Zhu, B., Zheng, X., Liu, H., Li, J. & Wang, P. Analysis of spatiotemporal characteristics of big data on social media sentiment with COVID-19 epidemic topics. *Chaos, Solitons & Fractals* **140**, 110123 (2020).
- [12] Gong, X., Han, Y., Hou, M. & Guo, R. Online public attention during the early days of the COVID-19 pandemic: Infoveillance study based on Baidu Index. *JMIR Public Health and Surveillance* **6**, e23098 (2020).
- [13] Shi, W.-z. *et al.* Online public opinion during the first epidemic wave of COVID-19 in China based on Weibo data. *Humanities and Social Sciences Communications* **9**, 1–11 (2022).
- [14] Cha, M., Haddadi, H., Benevenuto, F. & Gummadi, K. P. Measuring user influence in Twitter: The million follower fallacy. *ICWSM* **10**, 10–17 (2010).
- [15] Bakshy, E., Hofman, J. M., Mason, W. A. & Watts, D. J. Everyone’s an influencer: quantifying influence on Twitter. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 65–74 (2011).

- [16] Almaslukh, A. *et al.* Top-k sentiment analysis over spatio-temporal data. *PeerJ Computer Science* **10**, e2297 (2024).
- [17] Hanny, D. & Resch, B. 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).