Digital Hate Speech in the Thailand-Cambodia Border Dispute: Authentic vs. Inauthentic Account Patterns

A Study of Social Media Behavior During the 2025 Border Conflict and Swedish Arms Sales Controversy

Moses Ngeth
iam@mosesngeth.com

Table of Contents

Table

Digital Hate Speech in the Thailand-Cambodia Border Dispute Over the Gripen: Authentic vs. Inauthentic Account Patterns

A Study of Social Media Behavior During the 2025 Border Conflict and Swedish Arms Sales Controversy

1. Executive Summary

This study collected 4,337 comments from 2,474 Facebook profiles across 13 pages during the 2025 Thailand-Cambodia border conflict and Swedish arms sales controversy. The research analyzed 237 profiles to examine who participated, what they posted, and whether inauthentic accounts behaved differently than authentic accounts.

Key Findings

- **Most accounts showed suspicious behavior.** Only a minority appeared clearly authentic. The majority showed indicators of manipulation or questionable authenticity.
- **Hate speech dominated discussions.** More than half of participants posted content containing dehumanization, ethnic attacks, or violence threats.
- **Country targeting showed clear patterns.** Cambodia, which was attacked, received the most hate speech. Sweden, which sold weapons, faced moderate targeting. Thailand, which used military force, received the least criticism. Online discussions concentrated attacks on the conflict victim rather than countries involved in military actions or policy decisions.
- **Accounts with multiple suspicious indicators (3-4 indicators) posted more hate speech.** These accounts posted problematic content at higher rates than authentic accounts. They focused more intensely on Cambodia and attacked multiple targets simultaneously.
- **Targeting countries strongly correlated with hate speech.** Comments attacking specific countries were highly problematic. Policy-focused criticism remained mostly neutral. Criticizing policies stayed constructive; attacking countries turned toxic.

Research Contributions

This study documents specific patterns of hate speech and suspicious account behavior during an arms sales controversy amid armed conflict. The five-indicator detection method

successfully identified accounts that behaved differently from authentic users. The research establishes baseline measurements for understanding how online discussions operate during international crises.

2. Introduction

Social media changes how people understand world events. When countries disagree, online conversations shape what people think. This study asks a simple question: How do you tell the difference between authentic (real) accounts and inauthentic (fake) accounts during controversies over arms sales amid armed conflicts? The research examines Facebook discussions in August 2025 during the controversy over Sweden's sale of Gripen fighter jets to Thailand amid the Thailand-Cambodia border conflict.

Sweden's decision to sell advanced fighter aircraft[1] to Thailand sparked intense debate on Facebook three countries. During the July 2025 border clash, Thailand used F-16 fighter jets[2] to attack Cambodian territory. Despite this military action against Cambodia, Sweden signed an agreement in August 2025 to sell four additional Gripen jets worth $556 million to Thailand. This created a three-way controversy involving the arms seller (Sweden), the buyer (Thailand), and the conflict victim (Cambodia).

The study collected 4,337 comments from 2,474 Facebook users across 13 different Facebook pages. These pages included Thai news outlets, Cambodian media, and the Swedish Embassy in Bangkok, representing all three countries involved in the controversy. Most posts happened between August 25-28, 2025.

The research developed a five-indicator detection method to identify authentic versus inauthentic Facebook accounts. The study analyzed 237 Facebook profiles in detail. Authentic (real) accounts and inauthentic (fake) accounts act differently online. The analysis found clear patterns in how Thai, Cambodian, and Swedish Facebook users discussed the Gripen sales controversy.

[1] Reuters. (2025, August 25). *Sweden agrees to sell four Saab Gripen fighter jets to Thailand*. https://www.reuters.com/business/aerospace-defense/sweden-agrees-sell-four-saab-gripen-fighter-jets-thailand-2025-08-25/

[2] Camboja staff (2025, July 24). *Thai, Cambodian Troops Clash in First Border Fighting Since May*. Cambojanews. https://cambojanews.com/thai-cambodian-troops-clash-in-first-border-fighting-since-may/

3. Background

Thailand and Cambodia have disputed their border for decades. In May 2025, military clashes erupted along the border, one Cambodian solider was shot kill by Thai[3]. Thailand deployed F-16 fighter jets and struck Cambodian territory. The attacks killed dozens[4] of people and damaged civilian infrastructure.

On August 25, 2025, Sweden signed an agreement to sell four Gripen fighter jets to Thailand for $556 million. The timing triggered criticism from observers across the region. Thailand had just attacked Cambodia using American-made F-16s. Sweden then agreed to provide additional advanced fighter aircraft to Thailand.

The decision appeared contradictory to Sweden's stated values. For decades, Sweden has seen as a champion of human rights and democracy[5]. Sweden had supported[6] Cambodia with development aid and democracy programs. In 2021, Sweden closed[7] its embassy in Cambodia, citing concerns about human rights and democratic decline. Yet in 2025, Sweden sold advanced weapons to Thailand, a country that had just used military force against its neighbor.

This created controversy involving all three countries, though with different levels of engagement. The most intense exchanges occurred between Cambodian and Thai social media users. Many Cambodians expressed concern about Sweden's decision[8]. Thai officials and supporters defended the purchase as legitimate self-defense[9]. Some Swedish

---

[3] (2025, May 28). *Cambodian soldier killed in clash with Thai army*. DW. https://www.dw.com/en/cambodian-soldier-killed-in-clash-with-thai-army/a-72700404?maca=en-rss-en-world-4025-rdf

[4] CamboJA Staff (2025, July 24). *Death Toll Rises as Border Clash Escalates*. Camboja. https://cambojanews.com/death-toll-rises-as-border-clash-escalates/

[5] Freedom House (2024, February 28). *Freedom World*. Freedomhouse. https://freedomhouse.org/country/sweden/freedom-world/2024

[6] Andersen, H., Larsson, K. A., & Öjendal, J. (n.d.). *SUPPORTING STATE-BUILDING FOR DEMOCRATISATION? A STUDY OF 20 YEARS OF SWEDISH DEMOCRACY AID TO CAMBODIA*. Eba.se. https://eba.se/wp-content/uploads/2019/06/2019-03-Swedish-Democracy-Aid-to-Cambodia-webb-komprimerad.pdf

[7] Sun, N. (2020, November 27). Swedish Embassy to Move to Bangkok As Cambodia Continues Democratic Decline. Khmer.Voanews.com. https://khmer.voanews.com/a/swedish-embassy-to-move-to-bangkok-as-cambodia-continues-democratic-decline/5678757.html

[8] Meng, S. (2025, September 7). *Petition Launched in Cambodia to Request Swedish-Thai Fighter Jet Deal Suspension*. Cambodianess.com. https://cambodianess.com/article/petition-launched-in-cambodia-to-request-swedish-thai-fighter-jet-deal-suspension

[9] The Government Public Relations Department (PRD) (2025, July 29). *Royal Thai Armed Forces Reaffirm the Right to Self-Defense Under International Law to Safeguard National Sovereignty and the Thai People*. Thailand.prd.go.th. https://thailand.prd.go.th/en/content/category/detail/id/3243/iid/409959

parliamentarians debated whether the arms sale contradicted Sweden's human rights principles[10]. Each country's media covered the controversy from different perspectives.

This study focused on Facebook because the platform hosts significant political discourse in the region[11]. The research collected data from pages in three languages: Thai, Khmer, and English. The controversy generated intense emotions and rapid information sharing. This environment created opportunities for both genuine public expression and potential manipulation through inauthentic accounts.

4. Research Method

This study analyzed 237 Facebook profiles to understand how authentic and inauthentic accounts spread hate speech during the August 2025 Thailand-Cambodia border dispute and Swedish arms sales controversy. Complete methodology details are provided in **Appendix A: Research Methodology**.

Data Collection

The study collected 4,337 comments from 2,474 unique Facebook profiles across 13 selected pages. These pages included Thai news outlets (Bangkok Post, The Nation), Cambodian media (Phnom Penh Post, Fresh News), individual content creators, and the Swedish Embassy in Bangkok. The study gathered data during August, 2025, the peak period following Sweden's Gripen fighter jet sale to Thailand. Pages were selected through keyword searches and comment volume assessment (100-1,000+ comments). Data collection faced geographic restrictions and tool limitations. Therefore, the total number of comments for each page is not the same as the total number of comments that were analyzed.

Sampling Method

The study used proportional stratified sampling to select 237[12] profiles. Each Facebook page contributed samples based on its size.

---

[10] (2025, August 25). *Freeze weapons sales to Thailand and other non-democratic countries*. Www.Dagensarena.se. https://www.dagensarena.se/opinion/freeze-weapons-sales-thailand-non-democratic-countries/

[11] Center for the Study of Organized Hate (2024, December 3). *Structural Changes in Information Ecosystem Essential to Combat Disinformation in Southeast Asia*. www.Csohate.org. https://www.csohate.org/2024/12/03/southeast-asia-disinformation/

[12] *Systematic sampling across 13 Facebook pages initially yielded 238 profiles. Data cleaning procedures identified one duplicate profile due to Facebook ID error, resulting in a final analytical sample of 237 unique profiles. This*

Facebook Page	Population	Sample Size	Sampling Rate
The Phnom Penh Post	342	35	10%
The Bangkok Post	292	29	10%
វណ្ណនេត ហ្វីត - Vannet Feed	290	29	10%
The Nation Thailand	251	25	10%
Embassy of Sweden in Bangkok	249	25	10%
A JA A-អាចារ្យ អេ	226	22	10%
Khaosod English	193	19	10%
Fresh News	182	18	10%
Khaosod - ข่าวสด	143	14	10%
Men Kimseng	124	12	10%
Cambodianess	85	8	9%
Tang Senghour	52	5	10%
Cambodian Human Rights Committee	45	4	9%
Total	**2,474**	**245**	**10%**

Authenticity Detection Methods

The study developed a five-indicator method to identify fake accounts. The research marked profiles as suspicious if they had: fewer than 100 friends, only shared content without creating original posts, received no engagement on their posts, displayed suspicious profile elements like AI photos, or appeared in known fake account databases. Profiles showing 0-1 indicators were classified as authentic. Those with 3-5 indicators were classified as inauthentic.

represents a 9.6% sampling rate from the total population of 2,474 profiles, with a margin of error of ±5.3% at 90% confidence level.

Content Analysis

The study classified comments as either problematic (containing hate speech) or neutral. The study then identified which country or entity each comment targeted: Cambodia, Thailand, Sweden, or military/weapons policies.

5. Findings: What the Data Reveals

5.1 Account Types and Content Patterns

This section presents the overall landscape of social media, Facebook, activity during the 2025 border conflict and arms sales controversy. The findings reveal patterns in who participated in online discussions and how they communicated.

Overall Sample Characteristics

The study analyzed 237 unique Facebook profiles that commented on posts about the Thailand-Cambodia border dispute and Swedish Gripen sales. These profiles came from 13 Facebook pages across three countries. Each profile contributed an average of 1.8 comments to the discussion.

Source Distribution Across Facebook Pages

The 237 profiles came from 13 Facebook pages across three countries. Four Thai pages provided 49.4% of the sample: Bangkok Post, The Nation Thailand, and Khaosod (Thai and English editions). Seven Cambodian pages contributed 42.2%: The Phnom Penh Post, Vannet Feed, AJA-A, Fresh News, Tang Kimsour, Cambodianess, and Cambodian Human Rights Committee. The Swedish Embassy in Bangkok represented diplomatic communication (10.5%). Men Kimseng, a former VOA broadcaster, provided independent commentary. Each page contributed profiles proportionally to its total comment volume, ensuring representative sampling across all sources.

Account Authenticity: A Mixed Population

The profiles showed varying levels of authenticity based on our five-indicator detection system:

Table 1: Account Authenticity Classification

Category	Number	Percentage
Authentic	94	39.66%
Suspicious	70	29.54%
Likely Inauthentic	16	6.75%
Inauthentic	2	0.84%
Locked Profile	53	22.36%
N/A (Disabled)	2	0.84%
Total	**237**	**100%**

Understanding the Locked Profiles

The 53 locked profiles deserve special attention. Manual investigation revealed these accounts displayed suspicious characteristics. They had fewer than 100 friends and used questionable names and profile photos. Facebook users lock their profiles to restrict access to their information. This behavior, combined with other suspicious indicators, suggests these accounts were likely inauthentic. The locked status prevented full verification.

Combined Authenticity Assessment

When study combines the clearly inauthentic categories with suspicious accounts, the picture becomes clearer:

- **Clearly authentic:** 94 profiles (39.66%)
- **Questionable or inauthentic:** 141 profiles (59.49%)
- **Unable to verify:** 2 profiles (0.84%)

This means approximately 60% of profiles showed some level of suspicious behavior. This is a significant finding. The majority of accounts participating in this online discussion raised authenticity concerns.

Hate Speech Prevalence: More Than Half Problematic

The study examined the content posted by these profiles. The results show problematic content dominated the conversation:

- **Problematic content:** 135 profiles (56.96%)
- **Neutral or clean content:** 102 profiles (43.04%)

More than half of all participants posted content containing offensive language, discrimination, harassment, or hate speech. This high prevalence indicates the border conflict and arms sales controversy generated intense negative sentiment online.

What This Tells Us

Three key patterns emerge from this overview:

First, authentic accounts represented less than half of all account. The online discussion was heavily influenced by accounts showing suspicious behavior patterns.

Second, problematic content outnumbered neutral content. The conversation was more hostile than constructive.

Third, the combination of questionable authenticity and high hate speech rates suggests the online environment was toxic. This raises serious questions about whether the discussion reflected genuine public sentiment or amplified manipulation.

The following sections examine these patterns in greater detail. The study analyze how authenticity levels relate to hate speech rates and which countries faced the most targeting.

5.2 Authenticity Breakdown: Real People vs. Fake Accounts

This section examines account authenticity patterns using our five-indicator detection method[13]. The findings show clear differences between authentic and suspicious accounts.

[13] *See Appendix A: Research Methodology (Section 3.4.3) for detailed indicator rationale.*

How the Five Indicators Performed

Each indicator revealed different aspects of suspicious behavior.

Table 2: Indicator Distribution Across All Profiles

Indicator	Yes	No	N/A (Lock Profile)	Total
1. Few Friends (less than 100)	34 (14.35%)	201 (84.81%)	2 (0.84%)	237
2. Only Sharing Content	70 (29.54%)	112 (47.26%)	55 (23.21%)	237
3. No Engagement	58 (24.47%)	124 (52.32%)	55 (23.21%)	237
4. Suspicious Profile	124 (52.32%)	111 (46.84%)	2 (0.84%)	237
5. Database Match	0 (0%)	235 (99.16%)	2 (0.84%)	237

Indicator 4 (Suspicious Profile) was most common. Over half of all profiles (124, 52.32%) showed suspicious characteristics like AI-generated photos, inconsistent information, or accounts with long posting gaps. This suggests many accounts were either compromised or created for manipulation.

Indicator 2 (Only Sharing) appeared in 70 profiles (29.54%). These accounts never created original content, only shared posts from others.

Indicator 3 (No Engagement) affected 58 profiles (24.47%). These accounts posted content that received zero likes or comments, indicating no genuine social connections.

Indicator 1 (Few Friends) was least common. Only 34 profiles (14.35%) had fewer than 100 friends. This may suggest that many inauthentic accounts participating in the Gripen controversy were created earlier and had time to accumulate friends.

Indicator 5 (Database Match) found zero matches. No profiles matched the researcher's database of accounts documented from 2023-2025 (primarily 2024) during online harassment campaigns in Cambodia. The database contains accounts that posted hate speech and repetitive comments, with many showing patterns suggesting automated account creation. The researcher collected these accounts to track suspicious behavior across controversies and identify whether the same accounts participate in multiple toxic campaigns. Zero matches occurred because suspicious accounts were likely created specifically for the Gripen controversy, or because different operators participated than in previous harassment campaigns.

Authentic Accounts (94 profiles, 39.66%)

These accounts scored 0 or 1 indicator. They typically had more than 100 friends, created original posts regularly, received engagement from followers, and used real profile photos. These accounts represented real people expressing genuine opinions.

Suspicious Accounts (70 profiles, 29.54%)

These accounts scored 2 indicators, falling into a gray area. Common patterns included suspicious profiles combined with sharing-only behavior, or suspicious profiles with no engagement.

Likely Inauthentic Accounts (16 profiles, 6.75%)

These accounts scored 3 or 4 indicators, showing multiple signs of manipulation. They typically combined suspicious profiles, sharing-only behavior, and no engagement.

Locked Profiles: Hidden Information

The 53 locked profiles (22.36%) prevented full investigation. The study could only assess three indicators: friend count, profile appearance, and database matches.

Of the 53 locked profiles:

- 9 profiles (16.98%) had fewer than 100 friends
- 27 profiles (50.94%) showed suspicious profile characteristics
- All 53 had no database matches

Half of locked profiles showed suspicious characteristics in visible indicators. However, we could not see their posting behavior or engagement levels. The study classified all locked profiles separately rather than labeling them inauthentic, maintaining research integrity.

What Section 4.2 Reveals

Four key findings emerge:

First, suspicious behavior was extremely common. Only 40% of accounts appeared clearly authentic.

Second, profile manipulation (Indicator 4) appeared most frequently. Over half of all accounts showed red flags in their profile appearance or history.

Third, our detection method successfully identified accounts with varying levels of suspicion. Higher indicator counts corresponded to more obviously suspicious accounts.

Fourth, locked profiles hid significant information. Half showed suspicious characteristics in visible indicators, but the study cannot know their full behavior.

The online discussion during the border conflict involved a substantial proportion of accounts showing suspicious characteristics. The next section examines what content these accounts posted and which countries they targeted.

5.3 Hate Speech Patterns and Country Targeting

This section examines which countries faced targeting and how often that targeting included problematic content. The findings reveal significant differences in who received attacks and the severity of those attacks. Each profile was classified based on the sentiment in their most relevant comment about the controversy.

Overall Targeting Distribution

The 237 profiles targeted different countries, but one country received far more attacks.

Table 3: Sentiment Target Distribution

Target Category	Number of Profiles	Percentage
Anti-Cambodia Sentiment	86	36.29%
Anti-Sweden Sentiment	35	14.77%
Multiple Targets	29	12.24%
Military/Weapons Criticism	27	11.39%
Anti-Thailand Sentiment	18	7.59%
Other/Neutral	42	17.72%
Total	**237**	**100%**

Cambodia faced the most targeting. More than one-third of all profiles (86, 36.29%) directed negative sentiment toward Cambodia, far exceeding any other country.

Thailand received the least targeting. Only 18 profiles (7.59%) directed negative sentiment toward Thailand, less than one-quarter of the anti-Cambodia targeting.

Sweden faced moderate targeting. 35 profiles (14.77%) criticized Sweden, primarily over the arms sales decision.

Multiple targets appeared frequently. 29 profiles (12.24%) posted comments that targeted more than one category. These comments spread negativity across multiple dimensions rather than focusing on a single target.

Problematic Content by Target

Not all negative sentiment used hateful language. Some criticism remained within acceptable bounds. Others crossed into hate speech, dehumanization, and threats.

Table 4: Problematic Content Rates by Target

Target Category	Total	Problematic	Problematic %	Neutral	Neutral %
Multiple Targets	29	24	82.76%	5	17.24%
Anti-Thailand	18	14	77.78%	4	22.22%
Anti-Cambodia	86	66	76.74%	20	23.26%
Anti-Sweden	35	24	68.57%	11	31.43%
Military/Weapons	27	6	22.22%	21	77.78%
Other/Neutral	42	1	2.38%	41	97.62%

Critical Discovery: Attacking Countries Equals Hate Speech

When people targeted specific countries, they were much more likely to use hateful language.

Table 5: Hate Speech Risk by Target Type

Target Type	Problematic Rate	Risk Level
Multiple Targets	82.76%	Highest Risk
Anti-Thailand	77.78%	Very High Risk
Anti-Cambodia	76.74%	Very High Risk
Anti-Sweden	68.57%	High Risk
Military/Weapons Policy	22.22%	Low Risk
Other/Neutral Topics	2.38%	Very Low Risk

What This Pattern Means:

High-Risk Categories (69-83% problematic): Comments targeting countries contained hate speech 70-80% of the time, with rates ranging from 69% (Anti-Sweden) to 83% (Multiple Targets).

Low-Risk Categories (2-22% problematic): Comments discussing policy issues or neutral topics remained mostly clean.

The Pattern is Clear: When comments targeted countries, they contained hate speech. When comments discussed policies, they remained neutral.

Types of Hate Speech: What the Attacks Looked Like

Analysis of the offensive keywords revealed distinct patterns in how different targets faced attacks. The keywords presented below include both original English comments and comments translated from Thai languages for analysis purposes.

Note: The offensive keywords and allegations in the following tables reflect statements made by commenters during the controversy. They do not represent findings or conclusions of this research. The study documents and analyzes what was said, not whether the claims are accurate.

Anti-Cambodia Attacks: Severe Dehumanization

Anti-Cambodia attacks showed the most severe and extensive hate speech patterns. The keywords below reveal nine distinct attack categories, from animal dehumanization to death wishes.

Table 6: Example Anti-Cambodia Offensive Keywords (Top Examples)

Attack Category	Representative Keywords
Dehumanization (Animals)	buffalo, dogs, pig, cattle, worse than dogs, dog brain buffalo intelligence, animals, barking
Country Name Mockery	Scambodia, Claimbodia, Scumbodians, Cumbodia
Economic Attacks	beggar country, poorest country, broke, no money, need to buy shoes first, peanut
Intelligence Insults	stupid, no brain, idiot, brainless, low IQ, uneducated, brain problems
Slavery References	slave, khmer slave, slave nation, slave forever, slave bastard
Historical Trauma	Pol Pot killed all smart ones, Khmer Rouge fertilizer, killed millions, genocide
Violence Threats	kill, shoot dead, destroy, eliminate, turn to dust, must become extinct
Scam Allegations	scammer, scam hubs, hub of deception, fraud, cheater
Existence Denial	shouldn't survive, nobody wants them, pathetic nation, hell is in Cambodia

Note: Comprehensive keyword list available in Annex Table A1

Dehumanization was pervasive. Attackers used animal comparisons extensively, stripping Cambodians of human dignity.

Country name mockery appeared frequently. Derogatory variations like "Scambodia" became shorthand for contempt.

Historical trauma exploitation appeared. Some comments weaponized the Khmer Rouge genocide with mocking language.

Violence threats and death wishes occurred. Many comments crossed from insults into direct threats.

Cambodia faced the most attacks. Eighty-six profiles (36%) targeted Cambodia with hate speech. The attacks included economic mockery, dehumanization, intelligence insults, historical trauma references, and violence threats.

Anti-Thailand Attacks: Theft and Invasion Allegations

Anti-Thailand content focused on different themes, though still highly problematic.

Table 7: Example Anti-Thailand Offensive Keywords (Top Examples)

Attack Category	Representative Keywords
Theft Allegations	thieves, thief country, Thievesland, Liersland, Siamese thieves, generation of thieves
Invasion Claims	invaded, invading, always invades, stealing land, violation
Dehumanization	lady boy, soi dog, animals, beasts, Thai animals, breed of bastard
Violence Wishes	all Thai die, shoot dead, family bombing threat, kill
Historical Grievances	oppressed, blood in hands, genocide, bloody hands, always been thieves
Democracy Attacks	fake democracy, military dictatorship, coup history, brutal regime, authoritarian
National Character	liar, lies, fake, nonsense, hypocrite, evil, dirty behavior, bad neighbor

Note: Comprehensive keyword list available in Annex Table A2

Theft allegations dominated: The theft theme appeared in nearly every anti-Thailand comment.

Invasion and aggression claims were central: Comments framed Thailand as an aggressive military power.

Violence wishes occurred: Some comments included explicit death wishes against Thai people.

Anti-Thailand attacks focused more on political allegations and historical grievances than the ethnic dehumanization seen in anti-Cambodia content.

Anti-Sweden Attacks: Hypocrisy and Blood Money

Anti-Sweden content took a different approach, focusing on moral criticism rather than ethnic attacks.

Table 7: Example Anti-Sweden Offensive Keywords (Top Examples)

Attack Category	Representative Keywords
Hypocrisy Allegations	hypocrite, fake human rights supporter, masked killer, two-faced, double standards
Money Motivations	money face, blood money, hungry for money, greedy, profit over morality, cheap
Democracy Contradictions	closed embassy Cambodia but open North Korea, democracy when convenient, selective
Weapon Responsibility	killing Cambodians, weapons to kill people, supporting aggression, war criminal
Curse and Karma	karma will punish, disasters, collapse, suffer, Russia will attack you
National Elimination	disappear from earth, dissolve, nuclear bombs, destroy completely
Moral Bankruptcy	no humanity, inhumane, devil country, evil, criminal, lack of compassion

Note: Comprehensive keyword list available in Annex Table A3

Hypocrisy allegations dominated: The core attack questioned Sweden's moral credibility.

Money motivations were central: Attackers accused Sweden of prioritizing profit over principles.

Curse and karma wishes appeared: Unlike other categories, anti-Sweden comments included supernatural threats.

Anti-Sweden attacks focused on moral failures and betrayal rather than ethnic characteristics.

Military/Weapons Criticism: Policy-Focused Discussion

Military and weapons criticism stood apart from country-specific attacks.

Table 8: Example Military/Weapons Keywords (Top Examples)

Keyword Type	Representative Keywords
Technical Terms (Neutral)	fighter jets, F-16, Gripen, bombs, missiles, combat aircraft, weapons, arms sales
Action Descriptions (Neutral)	bombing, attacking, dropping bombs, shooting, invasion, border conflict, strike
Threat Language (Problematic)	test on your people, express mail, boom boom, human shields, destroy, target practice
Policy Criticism (Neutral)	arms exports, military cooperation, defense agreement, security concerns, regional stability

Note: Comprehensive keyword list available in Annex Table A4

Technical language dominated: Comments used military terminology that remained descriptive rather than emotional.

Most remained policy-focused: The 77.78% neutral rate shows commenters could criticize weapons sales while remaining within acceptable bounds.

The low problematic rate shows that discussing military topics did not inherently produce hate speech.

Three Critical Patterns

The targeting data and keyword analysis reveal three key findings:

Pattern 1: Cambodia Faced Disproportionate Attacks

This disproportionate targeting alone raises questions. The next pattern makes it even more concerning:

- **36.29%** of all profiles targeted Cambodia
- **Only 7.59%** targeted Thailand
- Cambodia faced **nearly 5 times more attacks** than Thailand

This reversed expectations. Cambodia, which was attacked, faced more online hate speech than Thailand, which used heavily military force**.**

Pattern 2: Anti-Cambodia Hate Speech Was Uniquely Severe

Cambodia faced not just more attacks, but uniquely severe attacks. The third pattern offers hope:

- Hundreds of distinct dehumanizing terms
- Nine different attack categories (see Table 6)
- Extensive animal comparisons and economic humiliation
- Historical trauma exploitation with Khmer Rouge references
- Violence threats and existence denial

The volume and variety exceeded all other categories.

Pattern 3: Policy Discussion Avoided Hate Speech

Military and weapons criticism showed that controversial topics could be discussed constructively:

- **77.78%** of military criticism remained neutral
- Technical language dominated over emotional attacks
- People criticized arms sales without ethnic attacks

This proves hate speech was avoidable when discussing the conflict.

5.4 Authenticity and Hate Speech: Two Critical Patterns

This section addresses the central research question: Do inauthentic accounts behave differently than authentic accounts? The study examine two specific questions: whether inauthentic accounts spread more hate speech, and whether they target different countries.

Research Question 1: Do Inauthentic Accounts Spread More Hate Speech?

The study cross-tabulated account authenticity levels with speech status to determine if inauthentic accounts posted problematic content at different rates than authentic accounts.

Table 9: Hate Speech Rates by Authenticity Level

Authenticity Level	Total	Problematic	Problematic %	Neutral	Neutral %
Authentic	94	51	54.26%	43	45.74%
Suspicious	70	37	52.86%	33	47.14%
Likely Inauthentic	16	11	68.75%	5	31.25%
Inauthentic	2	2	100%	0	0%
Locked Profile	53	34	64.15%	19	35.85%
N/A	2	0	0%	2	100%
Total	**237**	**135**	**56.96%**	**102**	**43.04%**

Finding: A Clear Pattern Emerges Only for Highly Suspicious Accounts

The data reveals a surprising pattern. Accounts with 2 indicators (Suspicious) actually posted problematic content at slightly LOWER rates than authentic accounts:

- **Authentic accounts:** 54.26% problematic
- **Suspicious accounts (2 indicators):** 52.86% problematic

This 1.4 percentage point difference shows that mildly suspicious accounts behaved nearly identically to authentic accounts in terms of hate speech rates.

However, a clear escalation pattern appeared for accounts with more indicators:

- **Likely Inauthentic (3-4 indicators):** 68.75% problematic
- **Inauthentic (4 indicators):** 100% problematic

Accounts scoring 3 or more indicators posted problematic content at significantly higher rates than authentic accounts. The difference of 14.49 percentage points between authentic (54.26%) and likely inauthentic (68.75%) represents a meaningful increase in hate speech behavior.

Locked profiles also showed elevated rates. At 64.15% problematic content, locked profiles fell between suspicious and likely inauthentic accounts.

What This Pattern Means:

The number of indicators is important. Accounts with multiple suspicious indicators (3+) behaved much worse than accounts with fewer indicators. This pattern reveals three insights:

- First, one suspicious indicator alone does not reliably identify inauthentic accounts. Many authentic users have unusual profiles or low engagement for legitimate reasons. Accounts with one indicator behaved similarly to clearly authentic accounts.
- Second, multiple suspicious indicators together identify problematic accounts more accurately. Accounts with three or more indicators showed measurably different behavior. Their hate speech rates were significantly higher than authentic accounts.
- Third, accounts with the most indicators posted only hate speech. Inauthentic accounts with all five indicators posted problematic content 100% of the time, with no exceptions.

Research Question 2: Do Inauthentic Accounts Target Different Countries?

The study examined whether inauthentic accounts focused their attacks on specific countries more than authentic accounts.

Table 10: Sentiment Targeting by Authenticity Level

Authenticity Level	Anti-Cambodia	Anti-Thailand	Anti-Sweden	Military/Weapons	Multiple Targets	Other/Neutral	Total
Authentic	36 (38.3%)	5 (5.3%)	16 (17.0%)	12 (12.8%)	9 (9.6%)	16 (17.0%)	94
Suspicious	21 (30.0%)	8 (11.4%)	13 (18.6%)	6 (8.6%)	5 (7.1%)	17 (24.3%)	70
Likely Inauthentic	7 (43.75%)	1 (6.25%)	1 (6.25%)	2 (12.5%)	4 (25.0%)	1 (6.25%)	16
Inauthentic	1 (50.0%)	0 (0%)	0 (0%)	0 (0%)	1 (50.0%)	0 (0%)	2
Locked	20 (37.7%)	4 (7.5%)	5 (9.4%)	7 (13.2%)	10 (18.9%)	7 (13.2%)	53
N/A	1 (50.0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (50.0%)	2
Total	**86**	**18**	**35**	**27**	**29**	**42**	**237**

Finding: All Groups Targeted Cambodia Most, But Highly Suspicious Accounts Showed Stronger Focus

Cambodia received the most targeting across all authenticity levels. Every category of accounts directed their primary attention toward Cambodia.

However, the intensity of Cambodia targeting varied by authenticity level:

- **Authentic accounts:** 38.3% targeted Cambodia

- **Suspicious accounts:** 30.0% targeted Cambodia (LOWER than authentic)
- **Likely Inauthentic:** 43.75% targeted Cambodia (HIGHER than authentic)
- **Inauthentic:** 50% targeted Cambodia

Likely inauthentic accounts focused most intensely on Cambodia. At 43.75%, these accounts with 3-4 indicators showed the strongest Cambodia focus. This represents a 5.45 percentage point increase over authentic accounts.

A surprising finding: Suspicious accounts were LESS focused on Cambodia than authentic accounts. Accounts with 2 indicators targeted Cambodia at only 30.0%, nearly 8 percentage points lower than authentic accounts (38.3%). Instead, these suspicious accounts posted more neutral content (24.3% vs 17.0% for authentic).

This counterintuitive result suggests that accounts with moderate suspicion indicators may include many legitimate users who simply have unusual profile characteristics but engage more neutrally.

Multiple targeting increased with suspicion level. Highly suspicious accounts spread their attacks across more targets:

- **Authentic:** 9.6% multiple targets
- **Suspicious:** 7.1% multiple targets
- **Likely Inauthentic:** 25.0% multiple targets
- **Inauthentic:** 50.0% multiple targets

Targeting Pattern by Authenticity Level:

Table 11: Key Targeting Characteristics by Authenticity Level

Authenticity Level	Cambodia Focus	Multiple Targets	Neutral Content	Key Behavior Pattern
Authentic (94)	38.3%	9.6%	17.0%	Moderate Cambodia focus with balanced distribution across categories
Suspicious (70)	30.0%	7.1%	24.3%	Lower Cambodia focus, higher neutral content, behavior similar to authentic accounts

Likely Inauthentic (16)	43.75%	25.0%	6.25%	Highest Cambodia focus, high multiple targeting, very low neutral content, clear shift toward aggressive targeting
Inauthentic (2)	50.0%	50.0%	0%	Split between Cambodia and multiple targets, zero neutral content, exclusively negative targeting
Locked (53)	37.7%	18.9%	13.2%	Cambodia focus similar to authentic, elevated multiple targeting, suggests mixed population

Note: Small sample sizes for Inauthentic (n=2) limit generalizability

Two key findings emerge from this examination of authenticity and hate speech patterns:

First, highly suspicious accounts (3+ indicators) showed clearly different behavior. These accounts posted problematic content at higher rates (68.75% vs 54.26% for authentic) and focused more intensely on Cambodia (43.75% vs 38.3%). The threshold effect matters. Multiple red flags together predicted worse behavior. Accounts with just one or two suspicious characteristics behaved similarly to authentic accounts, but accounts with three or more indicators showed a dramatic shift toward hate speech and aggressive targeting.

Second, all authenticity levels targeted Cambodia disproportionately, but highly suspicious accounts intensified this pattern. Every group of accounts directed their primary attention toward Cambodia, the victim in the border conflict. However, likely inauthentic accounts showed the strongest Cambodia focus and the highest rates of attacking multiple targets simultaneously. This suggests these accounts aimed to inflame tensions broadly rather than engage in balanced discussion.

These findings provide clear evidence that highly suspicious accounts behaved worse than authentic accounts. The 5-indicator detection system successfully identified accounts that posted more hate speech and focused more aggressively on vulnerable targets. While we cannot determine from this data whether these accounts operated independently or as part of coordinated efforts, their behavioral patterns differ measurably from authentic accounts.

In Summary for the findings

Finding 1: Most online participants showed suspicious behavior. Only 40% of accounts appeared clearly authentic. The remaining 60% displayed at least some indicators of manipulation or questionable authenticity. This suggests online discussions during conflicts are heavily influenced by accounts that may not represent genuine public sentiment.

Finding 2: Hate speech dominated the conversation. More than half (56.96%) of all participants posted problematic content including dehumanization, ethnic attacks, and violence threats. This toxic environment makes constructive dialogue nearly impossible.

Finding 3: The conflict victim faced the most attacks. Cambodia, despite being invaded, received nearly five times more targeting than Thailand, the aggressor. This reversed pattern suggests online discussions did not reflect the actual conflict dynamics. Anti-Cambodia hate speech was uniquely severe, employing hundreds of distinct dehumanizing terms across nine attack categories.

Finding 4: Highly suspicious accounts behaved measurably worse. Accounts with three or more manipulation indicators posted hate speech at rates 14 percentage points higher than authentic accounts (68.75% vs 54.26%). They focused more intensely on Cambodia (43.75% vs 38.3%) and spread attacks across multiple targets (25% vs 9.6%). This proves that inauthentic accounts amplified toxicity beyond what genuine users produced.

These patterns have serious implications. Social media platforms provided channels for severe hate speech targeting a conflict victim. Inauthentic accounts amplified this toxicity. The online environment during the crisis was poisoned by a combination of authentic users posting hateful content and suspicious accounts intensifying those attacks. Understanding these dynamics is essential for protecting democratic discourse during international conflicts.

6. Conclusion

This study examined Facebook discussions during the August 2025 controversy over Sweden's sale of Gripen fighter jets to Thailand amid the Thailand-Cambodia border conflict. The findings reveal how hate speech and suspicious accounts operated during this border conflict and arms sales controversy.

The Main Finding: Online Discussions Showed a Distorted Picture

The research documented three critical patterns during the Gripen sales controversy.

Hate Speech Dominated Discussions. More than half of participants posted content containing dehumanization, ethnic attacks, and violence threats. Policy criticism remained mostly neutral and constructive. However, when participants targeted specific countries, the discussion turned toxic. This shows hate speech was not inevitable during the controversy. The choice to attack countries rather than critique policies drove the toxicity.

Country Targeting Showed Clear Patterns. The research analyzed which countries received attacks during the controversy. Cambodia faced the most targeting, followed by Sweden, then Thailand. This distribution reveals a significant problem. Cambodia, which was attacked, received the highest proportion of hate speech. Thailand, which used military force, received the lowest proportion. Sweden, which sold the weapons, fell in between. The pattern shows that online discussions targeted the conflict victim most intensely, not the countries involved in military action or arms sales decisions.

Suspicious Accounts Amplified The Problem. Most online participants showed suspicious behavior patterns. Only a minority of accounts appeared clearly authentic. Accounts displaying multiple suspicious characteristics posted hate speech at higher rates than authentic accounts. These suspicious accounts also concentrated their attacks more intensely on Cambodia and targeted multiple countries simultaneously. The data reveals clear behavioral differences between authentic and suspicious accounts. Suspicious accounts made the toxic environment worse.

Why This Matters

These patterns have serious consequences for democracy. When inauthentic accounts manipulate discussions during conflicts, they distort public understanding of events. Online conversations showed the opposite of what actually happened on the ground. This distortion risks influencing public opinion and diplomatic relations in dangerous ways. People form

conclusions based on artificial online sentiment rather than facts. Democratic discourse suffers when citizens cannot distinguish genuine voices from manufactured outrage.

Implications for Stakeholders

- Social media platforms provided channels for severe hate speech targeting conflict victims without adequate safeguards. Platform algorithms amplified toxic content during a critical period when accurate information mattered most. The failure to detect and limit inauthentic account behavior allowed manipulation to flourish.
- Governments and civil society face challenges in protecting citizens from coordinated online attacks during conflicts. Traditional diplomatic responses cannot address the speed and scale of digital manipulation. Regional organizations need new approaches to monitor and counter artificial amplification of tensions.
- Citizens and media consumers lack tools to recognize when discussions are being manipulated. Media literacy efforts have not kept pace with the sophistication of inauthentic operations.

A Path Forward

- This research demonstrates that suspicious accounts can be identified through observable behavioral patterns. The five-indicator detection method successfully distinguished accounts that behaved differently from authentic users. This creates opportunities for intervention.
- Platforms can deploy similar detection systems to identify and limit suspicious accounts during conflicts. Governments can coordinate regional responses to cross-border hate campaigns. Civil society organizations can educate citizens to recognize manipulation tactics. These actions require commitment but remain achievable.
- The findings establish baseline patterns for measuring progress. Future research could consider this methodology when examining subsequent conflicts to assess whether interventions reduce hate speech and improve account authenticity. Understanding how inauthentic accounts operate is the first step toward protecting democratic discourse during the international crises that will inevitably come.

7. Recommendations:

For Social Media Users

1. **Criticize policies, not people.** You can criticize government decisions, military actions, and arms sales policies. Do not attack people with hate speech or

dehumanizing language. Policy criticism stays constructive. Personal attacks turn toxic.

2. **Learn to recognize and report hate speech.** Understand what counts as hate speech: dehumanization, ethnic attacks, violence threats, and calls for harm. When you see this content, report it to Facebook immediately. Do not share it. Do not engage with it. Reporting helps platforms identify patterns.
3. **Watch for inauthentic accounts.** Check profiles before sharing their content. Look for warning signs: fewer than 100 friends, only shares content without original posts, no engagement on their posts, suspicious profile photos, or newly created accounts. Do not let fake accounts manipulate you during conflicts. Do not help them spread hate speech.
4. **Wait before sharing emotional content.** During controversies, verify information through credible news sources before sharing.

For Regional Digital Rights Organizations

1. **Establish regional monitoring networks.** Create monitoring teams across the region. Share data about hate speech patterns and inauthentic account activity during border conflicts and arms sales controversies. Coordinated monitoring produces better evidence than isolated efforts.
2. **Work with social media companies to improve detection systems.** Partner with Meta, and other platforms to develop better hate speech detection for Southeast Asian languages. Share your research findings. Provide cultural context. Help platforms understand how hate speech works in Khmer, Thai, and other regional languages.
3. **Document and preserve evidence.** Archive hate speech content and inauthentic account activity before platforms delete it. Build secure databases for research, and legal accountability. Document evidence with timestamps and profile information.
4. **Educate the public.** Create simple educational materials in local languages showing how to recognize inauthentic accounts. Media literacy reduces manipulation effectiveness.
5. **Support victims of coordinated attacks.** Set up support systems for individuals and communities receiving death threats or coordinated harassment. Connect victims to legal resources, mental health support, and technical assistance for securing their accounts.

For Meta (Facebook)

This research did not analyze Meta's internal detection systems or content moderation processes. The study examined publicly visible hate speech that remained on the platform during the Gripen sales controversy. The presence of this content provides evidence that current detection and moderation systems did not prevent or quickly remove hate speech during this period. The recommendations below are based on observed outcomes: 57% of analyzed comments contained hate speech, hundreds of dehumanizing terms circulated publicly, and accounts with multiple suspicious indicators operated openly. These observations suggest opportunities for improvement.

1. **Hire native Khmer and Thai content moderators.** English-language AI systems miss hate speech in Southeast Asian languages. Recruit native speakers who understand cultural context, political tensions, and regional conflicts.
2. **Update hate speech detection for Khmer and Thai languages.** Current systems miss numerous of dehumanizing terms documented in this research. Review the comprehensive keyword lists in this study. Work with regional digital rights organizations and local experts to improve detection accuracy.
3. **Work with trusted partners to remove violating content and profiles.** Collaborate with regional digital rights organizations, local civil society groups, and independent researchers to identify hate speech and inauthentic accounts. Establish clear reporting channels. Remove identified violating content and profiles promptly.
4. **Adjust algorithms during arms sales controversies and border conflicts.** Stop amplifying content with ethnic slurs, dehumanization, or country mockery during active conflicts. Reduce the reach of posts containing violence threats or calls for harm. Crisis-period algorithms should prioritize accuracy over engagement.

References

Andersen, H., Larsson, K. A., & Öjendal, J. (n.d.). *Supporting state-building for democratisation? A study of 20 years of Swedish democracy aid to Cambodia*. Expertgruppen för biståndsanalys (EBA). https://eba.se/wp-content/uploads/2019/06/2019-03-Swedish-Democracy-Aid-to-Cambodia-webb-komprimerad.pdf

CamboJA Staff. (2025, July 24). Death toll rises as border clash escalates. *Cambojanews*. https://cambojanews.com/death-toll-rises-as-border-clash-escalates/

CamboJA Staff. (2025, July 24). Thai, Cambodian troops clash in first border fighting since May. *Cambojanews*. https://cambojanews.com/thai-cambodian-troops-clash-in-first-border-fighting-since-may/

Center for the Study of Organized Hate. (2024, December 3). *Structural changes in information ecosystem essential to combat disinformation in Southeast Asia*. https://www.csohate.org/2024/12/03/southeast-asia-disinformation/

Freeze weapons sales to Thailand and other non-democratic countries. (2025, August 25). Dagens Arena. https://www.dagensarena.se/opinion/freeze-weapons-sales-thailand-non-democratic-countries/

Freedom House. (2024, February 28). *Freedom in the world: Sweden*. https://freedomhouse.org/country/sweden/freedom-world/2024

Cambodian soldier killed in clash with Thai army. (2025, May 28). DW. https://www.dw.com/en/cambodian-soldier-killed-in-clash-with-thai-army/a-72700404?maca=en-rss-en-world-4025-rdf

Meng, S. (2025, September 7). Petition launched in Cambodia to request Swedish-Thai fighter jet deal suspension. *Cambodianess*. https://cambodianess.com/article/petition-launched-in-cambodia-to-request-swedish-thai-fighter-jet-deal-suspension

Reuters. (2025, August 25). *Sweden agrees to sell four Saab Gripen fighter jets to Thailand*. https://www.reuters.com/business/aerospace-defense/sweden-agrees-sell-four-saab-gripen-fighter-jets-thailand-2025-08-25/

Royal Thai Armed Forces. (2025, July 29). *Royal Thai Armed Forces reaffirm the right to self-defense under international law to safeguard national sovereignty and the Thai people*. Thailand Government Public Relations Department. https://thailand.prd.go.th/en/content/category/detail/id/3243/iid/409959

Sun, N. (2020, November 27). Swedish embassy to move to Bangkok as Cambodia continues democratic decline. *Voice of America Khmer*. https://khmer.voanews.com/a/swedish-embassy-to-move-to-bangkok-as-cambodia-continues-democratic-decline/5678757.html

Appendices

Appendix A: Detailed Methodology

Appendix B: Profile Analysis (Request only)