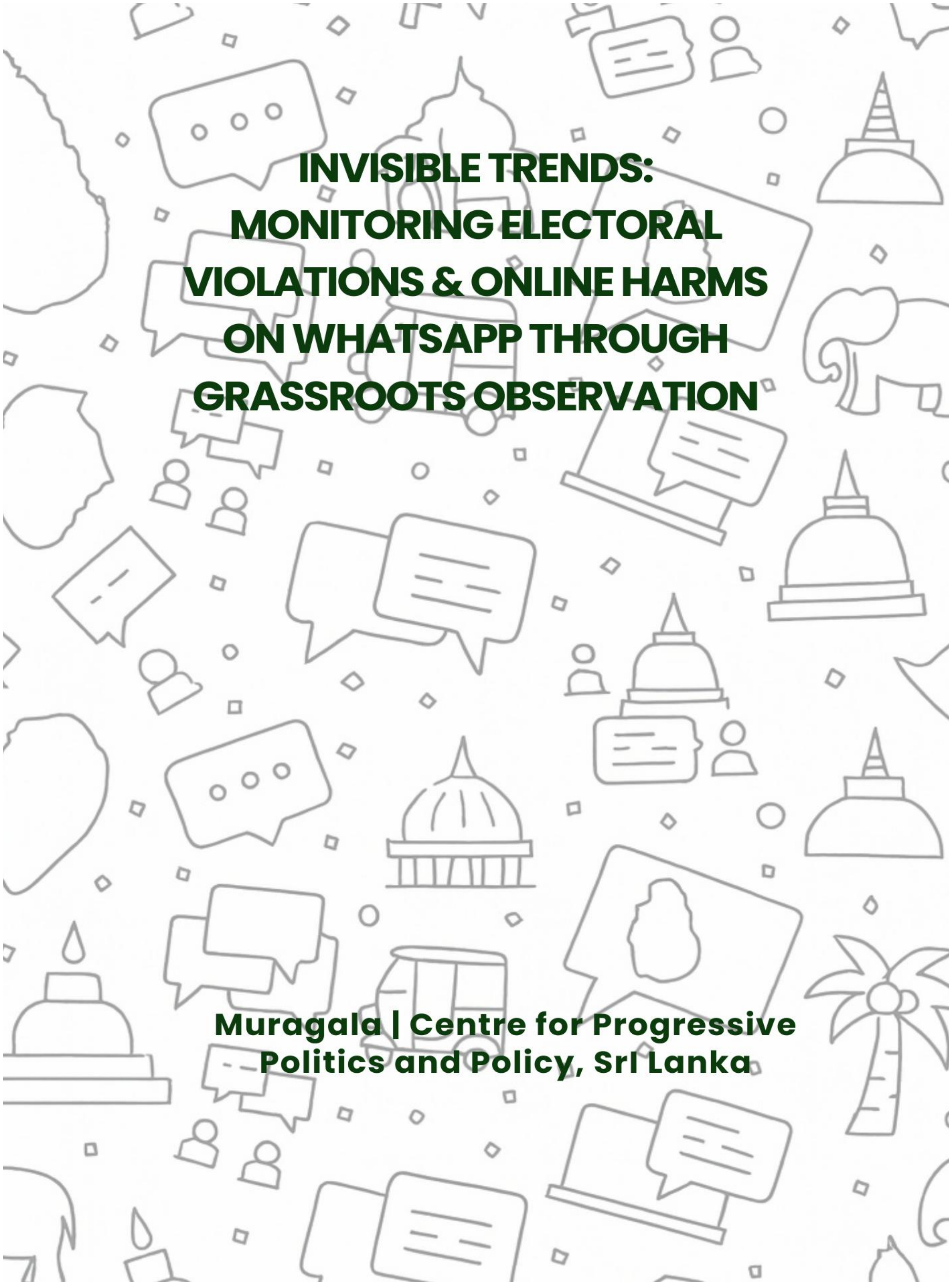
Co-funded by
the European Union

MURAGALA | CPPP
CENTRE FOR PROGRESSIVE POLITICS & POLICY



INVISIBLE TRENDS: MONITORING ELECTORAL VIOLATIONS & ONLINE HARMS ON WHATSAPP THROUGH GRASSROOTS OBSERVATION

Muragala | Centre for Progressive
Politics and Policy, Srl Lanka

`

Invisible Trends:
Monitoring Electoral Violations & Online Harms on WhatsApp through Grassroots Observation

December 2025

MURAGALA | CPPP
CENTRE FOR PROGRESSIVE POLITICS & POLICY

`

Table of contents

1. Introduction

1.1. WhatsApp and election campaigns

1.2. Monitoring WhatsApp in Sri Lanka’s 2024/25 election cycle

1.3. Research Focus on WhatsApp in Sri Lanka’s Elections

2. Methodology

2.1. Data Collection

2.2. Time Period

The data covered election-related communication between 10 April 2025 and 10 June 2025. This timeframe included the lead-up to the election in early May as well as the post-election period, extending until the establishment of Local Authorities (LGs), which was expected to be completed by 2 June 2025. Monitoring election-related violations stopped on June 10, a week after the establishment of the LGs.

2.3. Accessing Data and Scope of Monitoring

2.4. Data Transfer and recording

2.5. Categorisation of content and data analysis

3. Findings

3.1. Frequency of electoral violations in Whatsapp Groups

3.2. WhatsApp Groups often circulate content from other platforms

3.3. Thematic Analysis

3.3.1. Major types of electoral violations

4. Conclusions and Recommendations

4.1. Key Observations

4.2. Recommendations

Abbreviations

ACMC – All Ceylon Makkal Congress

ICT – Information and Communication Technology

ITAK – Illankai Tamil Arasu Kachchi

LGs – Local Authorities

NDF – National Democratic Front

NPP – National People’s Power

SJB – Samagi Jana Balawegaya

SLMC – Sri Lanka Muslim Congress

SLPP – Sri Lanka Podujana Peramuna

TFGBV – Technology-Facilitated Gender-Based Violence

UNP – United National Party

X – The platform formerly known as Twitter

`

Executive Summary

- WhatsApp emerged as a central but opaque campaign infrastructure during Sri Lanka's 2025 Local Government elections, enabling highly coordinated political messaging through private, encrypted group networks that largely escaped public scrutiny and regulatory oversight.
- Electoral violations were limited in volume but highly concentrated, with approximately 1% of monitored messages constituting violations, while a small number of groups recorded violation rates as high as 30%, indicating that targeted monitoring and interventions can be effective.
- WhatsApp functioned primarily as an amplification channel, not an originator: over 90% of violating content originated on public platforms (mainly Facebook and YouTube) and was recycled into WhatsApp groups, where it spread more rapidly and with less contestation.
- Four dominant categories of harm were identified: (i) misinformation and disinformation, (ii) hate speech and defamation, (iii) technology-facilitated gender-based violence (TFGBV), and (iv) systematic violations of the election "silent period," with the latter forming the single largest category.
- Gendered and identity-based attacks were particularly pronounced, disproportionately targeting women politicians, LGBTQI+ individuals, and minority candidates through sexist slurs, sexualised defamation, and recycled smear campaigns with long shelf-lives.
- Coordinated "cascade" group structures enabled rapid, nationwide dissemination, often mirroring state administrative divisions, blurring boundaries between party communication and official authority and increasing the persuasive power of campaign messaging.
- Party presence did not always correlate with violation intensity: while larger parties had more groups, some parties showed disproportionately higher rates of violations, particularly in relation to hate speech and TFGBV, suggesting strategic rather than incidental misuse.
- Grassroots, observer-based monitoring proved both feasible and effective, generating granular, real-time insights into encrypted environments that are otherwise inaccessible to regulators, researchers, and platforms.
- Election regulators should formally integrate WhatsApp into election-monitoring frameworks, including silent-period enforcement, guidance to political parties, and collaboration with civil society monitors to identify coordinated abuses.
- Meta/WhatsApp should be urged to adopt election-specific safeguards, including clearer political content policies for groups and channels, transparency around bulk messaging and automation, and mechanisms that enable privacy-respecting oversight during election periods.

`

1. Introduction

1.1. WhatsApp and election campaigns

This report is an analysis of how WhatsApp was used during and after the Local Government elections in Sri Lanka held in May 2025. Despite the increasing dominance of WhatsApp as a key social engagement and communication platform, it remains significantly under-scrutinized in both academic and regulatory circles. Its accessibility makes it especially valuable in low-resource settings where traditional media is limited, and it has rapidly become a central platform for political discourse. Embedded in everyday life across many regions, WhatsApp serves for many citizens in democracies as a primary tool for coordination and information exchange. Its versatility, encompassing text and voice messaging, file sharing, and group management, has earned it the moniker of an "everything app."[1] As a result, the platform has become an ubiquitous and cost-effective mode of communication, widely used for interpersonal messaging, consuming news and entertainment, and coordinating daily activities.[2]

In Sri Lanka, with 63% of the population actively using WhatsApp as of late 2024, it is the most widely used social media platform in the nation.[3] This prominence is often attributed to low data consumption, mobile-first accessibility, and its multifunctional use for messaging, entertainment, and news sharing. The platform is currently the second largest in the world in terms of the number of subscribers, and has achieved market dominance in many large and small countries in the Global South, such as India, Brazil, Indonesia, and Sri Lanka. Because of its encrypted and viral messaging features, the app has been frequently exploited to spread political disinformation, incite polarization, and undermine electoral integrity, posing serious risks to democratic processes.[4]

[1] Lapowsky, I. (2024, December 22). *How WhatsApp became the world's 'everything app'*. Rest of the World. https://scroll.in/article/1076663/how-whatsapp-became-the-worlds-everything-app.

[2] Lee, C. E., Chern, H. H., & Azmir, D. A. (2023). WhatsApp Use in a Higher Education Learning Environment: Perspective of Students of a Malaysian Private University on Academic Performance and Team Effectiveness. *Education Sciences*, *13*(3), 244.

[3] Capital Alliance Limited. (2024, October). *WhatsApp is Sri Lanka's top social media platform, used by 63% of the population*. The Morning. https://www.themorning.lk/articles/BasNOlXTMzjjjPcvWA55.

[4] Hale, S. A., Belisario, A., Mostafa, A. N., & Camargo, C. (2024). *Analyzing misinformation claims during the 2022 Brazilian general election on WhatsApp, Twitter, and Kwai*. *International Journal of Public Opinion*

`

Research indicates that WhatsApp's private, encrypted, and interpersonal nature significantly contributes to users' increased susceptibility to harmful speech and reduced resistance to misinformation.[5] WhatsApp, unlike public platforms such as Facebook, X, TikTok, or YouTube, fosters closed networks of trust, often among family, friends, or colleagues, where messages are accepted with minimal scrutiny. The absence of public profiles and moderation mechanisms diminishes accountability, while end-to-end encryption creates a sense of impunity that encourages the circulation of provocative or extreme content. Repeated exposure to such speech within intimate settings can normalize harmful narratives, making them part of everyday discourse. Moreover, political actors have strategically exploited WhatsApp's virality and emotional intimacy to disseminate polarizing propaganda, particularly during elections, leveraging users' trust in their contacts to bypass critical evaluation and amplify divisive messages. According to Garimella & Chauchard (2024), "despite WhatsApp's ubiquity and its acknowledged role in spreading problematic content, research on WhatsApp remains relatively low, mainly due to the challenging nature of obtaining data for studying communication on the platform. This contrasts with platforms like Facebook and Twitter, where user data has been more accessible for academic analysis".[6]

However, a growing body of literature has looked at WhatsApp to understand what goes on inside this platform and how politically harmful materials can be addressed. One recent study on WhatsApp Groups and Channels showed how the platform lacks election misinformation policies that are enforced on Meta's other platforms.[7] While Facebook and Instagram have clearly articulated guardrails against voter interference and disinformation, WhatsApp Channels

Research, 36(3); Ashish, C. R. (2025). *Fake news and electoral democracy: A content analysis of WhatsApp usage in South India*. *International Journal of Creative Research Thoughts, 13*(6), 581; Udupa, S., & Wasserman, H. (2025). *WhatsApp in the world: Disinformation, encryption, and extreme speech*. NYU Press.

[5] Arnaudo, D. (2017). *Computational propaganda in Brazil: Social bots during elections* (Working Paper No. 2017.8). Oxford Internet Institute, University of Oxford. Retrieved from DemTech website: https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/06/Comprop-Brazil-1.pdf.

[6] Garimella & Chauchard (2024), WhatsApp Explorer: A Data Donation Tool to Facilitate Research on WhatsApp https://www.gla.ac.uk/media/Media_1153129_smxx.pdf.

[7] Kern, R. (2024, June 7). *WhatsApp Channels, used by millions, have no clear election rules.* Politico.https://www.politico.com/news/2024/06/07/meta-election-guardrails-whatsapp-channels-00161073.

`

and chat groups operate without explicit rules, allowing political content that might be removed elsewhere to circulate freely.

Suggesting a causal link between WhatsApp usage and exposure to misinformation, one group of Brazilian researchers performed a deactivation experiment, where users were invited to avoid using WhatsApp for the period of study. The study revealed that for many participants, the platform served as a significant channel for political misinformation during elections.[8] The study found that reducing WhatsApp exposure lowered recall of false rumours, suggesting that the platform plays a unique role in spreading content that might not thrive on traditional feeds. Another investigation comparing WhatsApp and Facebook found that WhatsApp's intimate, closed communication style facilitates the spread of misinformation.[9] Our study in Sri Lanka confirms this observation, illustrating how content that might be challenged or corrected on public platforms often goes unchecked in WhatsApp groups, where social corrections are less frequent and users are more trusting of senders.

Recently added features to WhatsApp have further complicated the platform's role. Despite being classified as a closed messaging app, the platform has several public broadcasting features such as Channels, which can have over 1000 members. This increasing ability to coordinate large groups of people, coupled with its lack of specific election misinformation policies, poses a significant threat to the integrity of communication on the platform. Further, the use of automated tools and bulk messaging services that violate WhatsApp's Terms of Service remain difficult to detect due to the platform's encrypted nature. These dynamics have prompted calls for more robust monitoring mechanisms that respect user privacy while ensuring electoral integrity.[10]

[8] Ventura, T., Majumdar, R., Nagler, J., & Tucker, J. A. (2025). *Misinformation beyond traditional feeds: Evidence from a WhatsApp deactivation experiment in Brazil*. The Journal of Politics. Advance online publication https://csmapnyu.org/research/academic-research/misinformation-beyond-traditional-feeds-evidence-from-a-whatsapp-deactivation-experiment-in-brazil.

[9] Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2021). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society, 23*(8), 2430–2451. https://doi.org/10.1177/1461444820928059

[10] Mozilla Foundation. (2024, April 2). WhatsApp: Reform features now to protect election integrity. https://www.mozillafoundation.org/en/blog/whatsapp-reform-features-now-to-protect-election-integrity/.

`

Although some scholarship has started exploring these trends, there is still substantial work to be done in assessing their accuracy and in unravelling the potential role of WhatsApp in electoral outcomes. Despite the messaging service's ubiquity, accessing WhatsApp data for research purposes remains a challenging task. While evidence about the "information diet" of Facebook and Twitter (now X) users has been available for some time now, systematic evidence of a similar nature is notably lacking for WhatsApp, even though the platform's likely role in spreading problematic content is often acknowledged.[11] In this context, our study seeks to understand how the platform was used for spreading election related misinformation and hate by looking at data from grassroots activists.

Grassroots-based monitoring has emerged as a promising approach to address these challenges. By embedding trained observers within local communities, researchers can document digital violations in real time, capturing context-specific patterns that are often missed by centralized oversight. This method enhances data granularity, supports ethical data collection, and empowers local actors to participate in safeguarding democratic processes.

Notable initiatives in Sri Lanka, Kenya, and the Philippines have demonstrated the efficacy of community-based monitoring in identifying problematic content, tracking misinformation flows, and informing policy recommendations.[12] These efforts underscore the need for interdisciplinary collaboration between civil society, academia, and technology platforms to develop scalable and context-sensitive monitoring frameworks.

[11] Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). *Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? Psychological Science*, 26(10), 1531–1542; Guess, A., Nagler, J., & Tucker, J. (2019). *Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook. Science Advances*, 5(1), https://doi.org/10.1126/sciadv.aau4586; Tucker, J. A., Guess, A. M., Barberá, P., Vaccari, C., Siegel, A. A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. https://dx.doi.org/10.2139/ssrn.3144139.

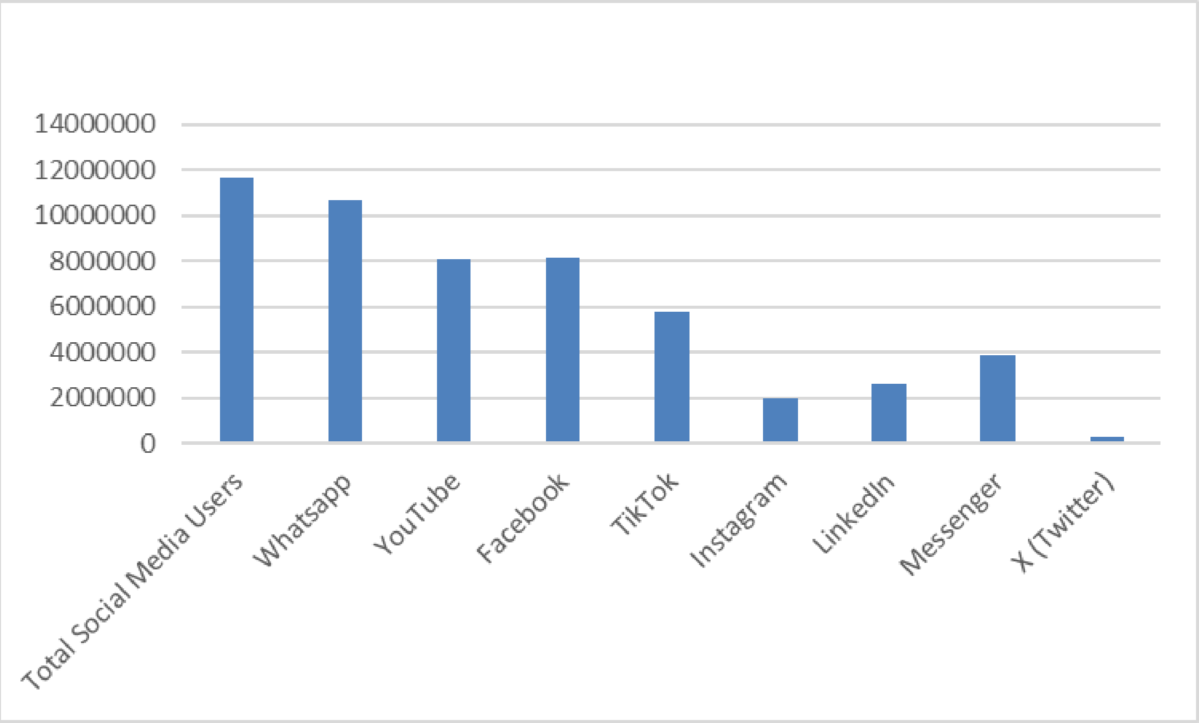
[12] LIRNEasia. (2024). *Election Misinformation in South and South-East Asia: The phenomenon and measures to counter it — Report draft*. IDRC; Mondini, R., Kotonya, N., Logan IV, R. L., Olson, E. M., Lungati, A. O., Odongo, D. D., Ombasa, T., Lamba, H., Cahill, A., & Tetreault, J. R. (2024). *Uchaguzi-2022: A Dataset of Citizen Reports on the 2022 Kenyan Election*. arXiv; Rappler & Internews. (2022). *Patient Zero: A study on the Philippine Information Ecosystem.* https://internews.org/wp-content/uploads/2022/02/Rappler-Internews-Patient-Zero-A-study-on-the-Philippine-information-ecosystem.pdf.

`

1.2. Monitoring WhatsApp in Sri Lanka’s 2024/25 election cycle

Between 2024 and 2025, Sri Lanka witnessed an intense electoral resurgence, holding three major elections within just eight months. This marked a dramatic return to democratic activity after a total absence of national polls since 2020. The compressed timeline reactivated political engagement across the country, creating fertile ground for new narratives and voter mobilization. During this period, WhatsApp penetration grew significantly alongside broader internet access. With mobile connectivity expanding rapidly, WhatsApp became one of the most accessible and widely used platforms for real-time communication. Its low data usage, group messaging features, and multimedia capabilities made it a preferred tool for both everyday interaction and political discourse.

Figure 1: Social Media User Count in Sri Lanka, 2025



(Sources: https://datareportal.com/reports/digital-2025-sri-lanka and https://gs.statcounter.com/social-media-stats/all/sri-lanka).

By 2025, with over 11 million users across all ages and regions, WhatsApp demonstrated its capacity to mobilize movements, organize dissent, and coordinate civic action.[13] The 2024–2025

[13] Editorial Staff. (2024, August 27). *Most Popular Social Media in Sri Lanka*. Tectera. Retrieved August 26, 2025, from https://tectera.com/most-popular-social-media-in-sri-lanka/

election cycle confirmed that social media was no longer peripheral to Sri Lanka's democracy. Given its cost-effective and decentralized nature, these platforms provided an open space for political expression, bypassing traditional gatekeeping institutions such as state media and regulatory bodies. They enabled citizens to challenge official narratives, amplify dissent, and organize around opposition movements.

WhatsApp, in particular, became a critical instrument for grassroots activism. Key figures from the National People's Power (NPP) acknowledged that their electoral success depended heavily on social media mobilization.[14] The NPP, which remained only the third-largest Sinhala-majority political force, adapted effectively to this digital shift, running an aggressive, tech-enabled campaign. Through a highly structured WhatsApp-based communication network, the party disseminated its messaging, engaged activists, and reached sympathizers directly. Other political parties, including the Samagi Jana Balawegaya (SJB), the National Democratic Front (NDF) and the United National Party (UNP), as well as several Tamil- and Muslim-led parties, also used WhatsApp in their campaigns, though generally not as effectively as the NPP. This digital-first strategy allowed the NPP to shape public opinion and secure decisive victories in all three elections, signaling a new era of social media-driven political mobilization in Sri Lanka.

WhatsApp's centrality to this shift stems from its increasing popularity among young adults and middle-aged users, who make up the largest share of Sri Lanka's social media landscape. As the platform most widely used for campaigning, community organizing, and rapid information sharing, it has become a core component of the country's digital public sphere. Political actors increasingly rely on WhatsApp's closed-group architecture to coordinate activities, disseminate tailored messages, and mobilise supporters across districts with minimal delay. Yet the same encrypted, privately networked structure that empowers grassroots coordination also complicates oversight, highlighting WhatsApp's dual character: an engine for civic participation, but also a potential conduit for misinformation and other digital harms during electoral cycles.

[14] Hattotuwa, Sanjana. 2024. "Social media, meditations, & mediations: Snapshots of Sri Lanka after the consequential elections in 2024." *Sanjana's Blog*. Published 30 November 2024, https://sanjanah.wordpress.com/2024/11/30/social-media-meditations-mediations-snapshots-of-sri-lanka-after-the-consequential-elections-in-2024/.

1.3. Research Focus on WhatsApp in Sri Lanka’s Elections

This study is guided by a central question: How can grassroots monitoring of WhatsApp contribute to safeguarding electoral integrity and mitigating digital harms? The inquiry is situated within Sri Lanka’s dynamic electoral environment, where WhatsApp has become both a tool for civic engagement and a channel for potential risks such as misinformation, hate speech, and coordinated manipulation.

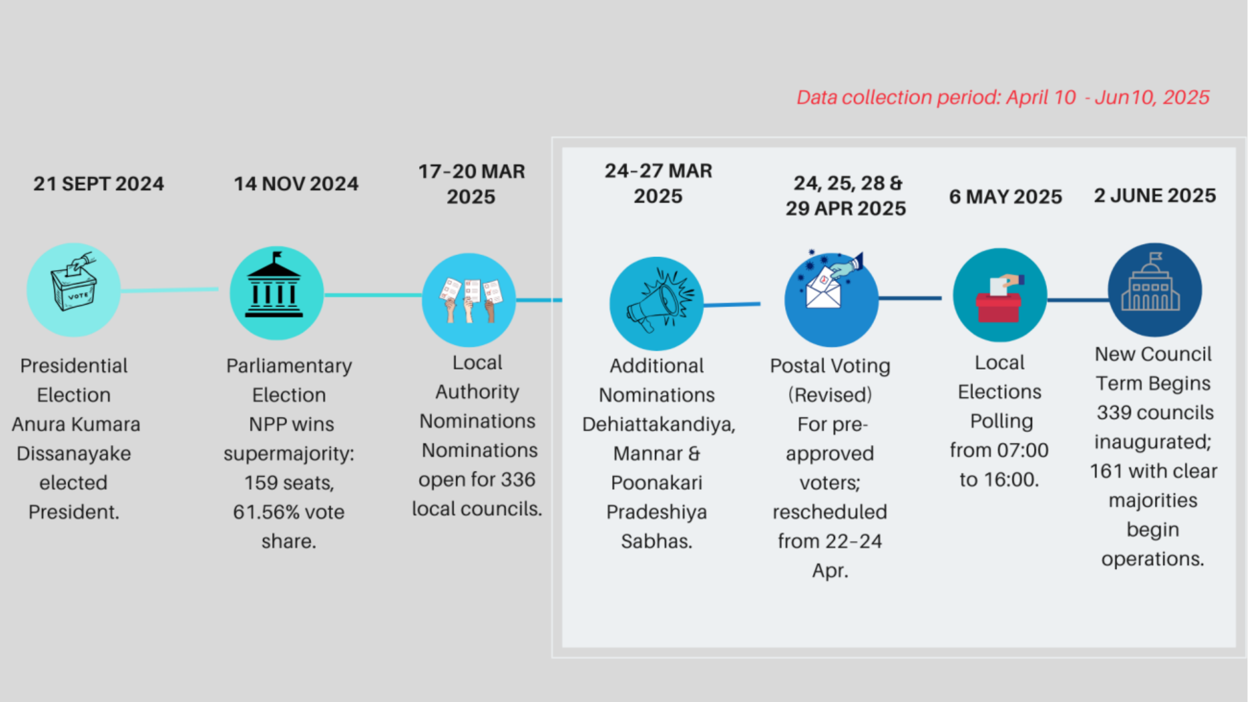
The research pursues to document the nature of digital violations on WhatsApp during elections, including misinformation, hate speech, OGBV and silent period violations, and coordinated campaigns. Based on these, the study seeks to develop policy recommendations for electoral safeguards and platform governance.

The scope of the study was defined by selecting geographic regions representing major electoral districts, including both Tamil- and Sinhala-speaking areas. A total of 30 locally based social activists participated in data collection, monitoring WhatsApp conversations within the groups they were already part of and could gain access to for the purpose of monitoring. To ensure broader exposure to WhatsApp-based political campaigns and potential violations, they were also encouraged to join a diverse range of groups and channels. The monitoring activities were carried out on a part-time basis between April 10 and June 10, 2025.

Ethical considerations, such as the linguistic and gender diversity of the monitoring volunteers, informed consent, and secure data handling, were central to the research design. Nonetheless, there are limitations, notably in the restricted access to encrypted content and potential observer biases in reporting. The natural limitation of this method is that only the content that appears in the subscribed WhatsApp groups are monitored. For instance, while all activists had access to some local political messaging group, such as by local candidates, local party activists or sympathisers, 7 candidates (23.3%) activists could not gain access to groups or channels of all major political parties or groups in their local area.

Figure 2: Election timeline in Sri Lanka 2024/ 2025

`



Data collection period: April 10 - Jun10, 2025

21 SEPT 2024
Presidential Election Anura Kumara Dissanayake elected President.

14 NOV 2024
Parliamentary Election NPP wins supermajority: 159 seats, 61.56% vote share.

17-20 MAR 2025
Local Authority Nominations Nominations open for 336 local councils.

24-27 MAR 2025
Additional Nominations Dehiattakandiya, Mannar & Poonakari Pradeshiya Sabhas.

24, 25, 28 & 29 APR 2025
Postal Voting (Revised) For pre-approved voters; rescheduled from 22-24 Apr.

6 MAY 2025
Local Elections Polling from 07:00 to 16:00.

2 JUNE 2025
New Council Term Begins 339 councils inaugurated; 161 with clear majorities begin operations.

This study examines harmful content circulating on WhatsApp during the LG election, with a focus on misinformation, silent-period violations, gendered hate speech, and online gender-based violence (OGBV).

It also checks if local, digital watch groups can help keep an eye on elections, working with standard monitoring. The goal is to give useful ideas to help support action and make institutions better at dealing with digital problems that happen around election time.

2. Methodology

Mobilising a grassroots-based monitoring approach, this study collected and analysed instances of messaging that was in violation of election laws, or was online political hate speech.

2.1. Data Collection

This study relied entirely on manual data collection. Monitoring activists acted as participants–observers in selected public WhatsApp groups. A group of 30 observers was deployed in all electoral districts in Sri Lanka. Using their personal devices and identities, and secondary accounts where necessary, activists joined a wide range of politically active groups to capture communication relevant to the election period.

2.2. Time Period

The data covered election-related communication between 10 April 2025 and 10 June 2025. This timeframe included the lead-up to the election in early May as well as the post-election period, extending until the establishment of Local Authorities (LGs), which was expected to be completed by 2 June 2025. Monitoring election-related violations stopped on June 10, a week after the establishment of the LGs.

2.3. Accessing Data and Scope of Monitoring

Activists gained entry to groups by using their personal or secondary accounts. The criteria for selection included whether the group was started by a political party or candidate, whether the discussions explicitly or implicitly related to electoral politics, and whether political communication intensified during the election cycle. Accordingly, 1743 groups and channels operated at local, regional, or national levels were observed. Those selected for monitoring were those that had been created or mobilized predominantly for political purposes. The criteria used for selecting groups or channels for monitoring included, the group's amplifying nature, language (the composition was Sinhala - 75% and Tamil - 25%) ,

Personal chats involving only two persons or small chat groups among close friends (such as family, schoolmates, or coworkers) were not considered for the study.

2.4. Data Transfer and Recording

When an abusive or potentially abusive message was identified, the content was immediately transferred to the coordinator's account. This ensured that suspected violations were captured in real time. All observed content was systematically recorded in a standardized monitoring template using a digital log sheet (Google Sheets). The codebook included fields for the observer's identity, details of the content producer or sharer, language, region or city, group size and origin, time and date stamp, and the text or screenshot of the content (with links provided for videos). Each entry also documented the nature of the violation, such as misinformation, hate speech, or online gender-based hate speech, along with a description of the content and translations where necessary. Additional notes captured related posts by the same user or observer comments.

`

2.5. Categorisation of Content and Data Analysis

Recorded content was categorized under three main types of violations:

i) Mis-, dis-, and mal-information targeting elections, including mock elections, discouraging voting (voter suppression), defined in terms of the Media Guidelines under Article 104B(5)(A) of the Constitution of Sri Lanka.

(ii) Hate speech, including targeting an ethnic group, gender, or disability.

(iii) Technology-facilitated gender-based violence (TFGBV): For a definition of this, WhatsApp's Terms of Service[15] was used in combination with the following definition below. An ad-hoc terminology was developed to categorise hate speech, and no machine assistance was used in the identification. To define hate speech, the following demarcation was used:

> "Online gender-based hate speech refers to digital content shared through ICT platforms that expresses hatred toward women (and girls) based on gender alone or in combination with other identity factors such as race, age, disability, sexuality, ethnicity, nationality, religion, or profession. It may involve spreading, inciting, promoting, or justifying gender-based hostility, and often includes violent or degrading material that portrays women (and girls) as sexual objects or targets of violence. Such content can circulate publicly or privately and is frequently directed at women in public-facing roles."[16]

Content analysis involved qualitative coding of messages and patterns, thematic analysis to identify types and trends of violations, quantitative summaries of the frequency of specific harms, cross-cutting analysis of gender-based targeting, and region- and language-based assessments. Network analysis maps how political parties and actors mobilize clusters of WhatsApp groups for campaigning and propaganda, particularly in the dissemination of harmful content

[15] WhatsApp. (n.d.). *How to use WhatsApp responsibly*. WhatsApp Help Center. Retrieved August 16, 2025, from https://faq.whatsapp.com/361005896189245

[16] European Institute for Gender Equality. (n.d.). *Safe Spaces: What is online gender-based hate speech?* Newsroom. Retrieved August 18, 2025, from https://eige.europa.eu/newsroom/safe-spaces/online-hate-speech?

`

The study prioritized ethical safeguards and safety of activists. Safety protocols were communicated in the training workshops. All monitors provided informed consent before participating in the study. Screenshots and other media were stored securely in encrypted folders. This study was not exhaustive or comprehensive. The monitoring activists had limited access to groups and channels, and the total number of relevant communication streams (groups) monitored were 1743. The activity of groups varied between 50 to 200 posts per day (discounting responses or reactions to posts).

3. Findings

3.1. Frequency of Electoral Violations in Whatsapp Groups

We first identified the types of violations that occur on WhatsApp. The study analyzed a total of 212,134 monitored messages, identifying 2,140 electoral violations in all the monitored groups. The monitoring exercise recorded 543 unique violations, as many posts were duplicated across several WhatsApp groups.

Figure 3: Percentage of violations

Total Messages Monitored: 212,134. Electoral violations: 2,140

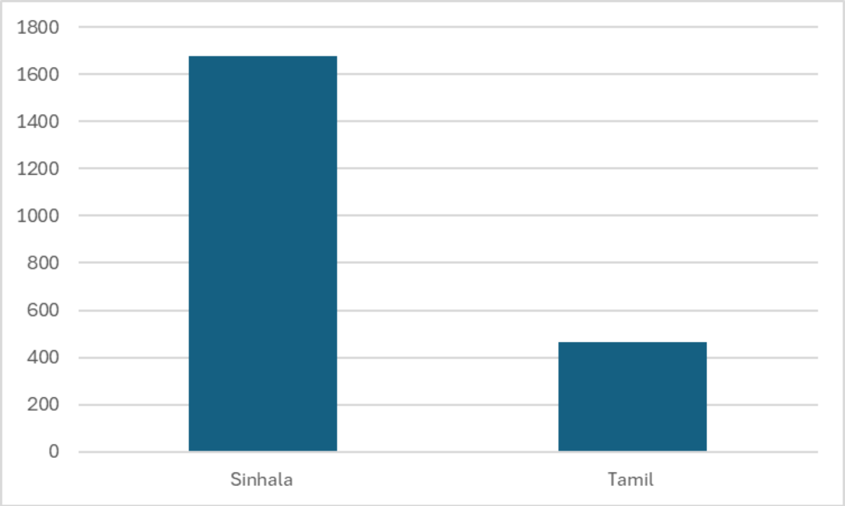
In some groups the rate may be as high as 30%. Others have negligible violations, after admin interventions.

This indi low, a significant discrepancy exists across different WhatsApp groups monitored. The violation rate in some groups was found to be as high as 30%, while others exhibited very low levels of violations, particularly following administrative interventions. This suggests that while violations are not widespread, they are highly concentrated within specific subgroups, highlighting the effectiveness of targeted interventions.

The language-wise breakdown of these violations show that Sinhala language-based groups have shown proportionately more frequent violations than Tamil language-based groups.

Figure 4: Electoral violations by language

`



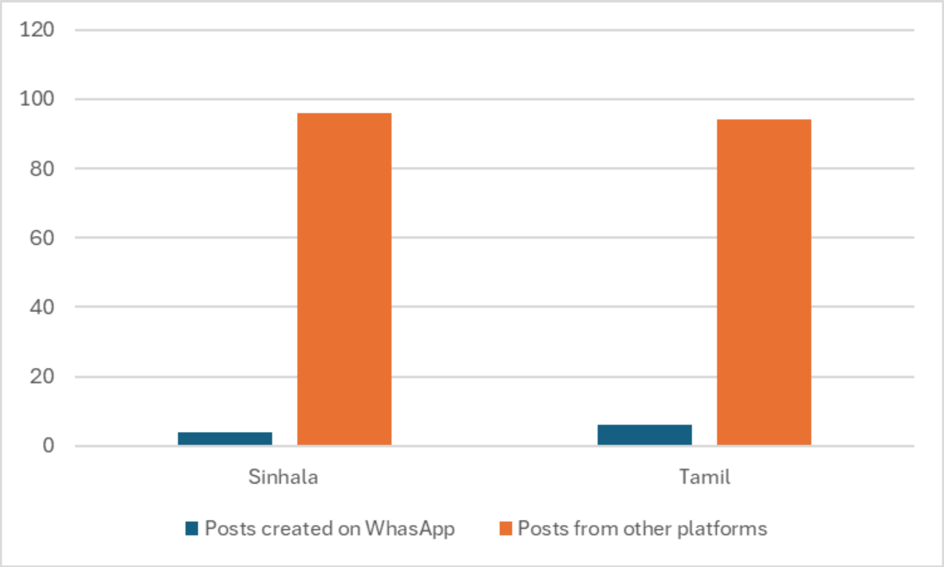
3.2. WhatsApp Groups Circulate Content from Other Platforms

The study seeks to understand how WhatsApp is leveraged to facilitate the amplification of electoral violations. We found that while WhatsApp was not just a platform for sharing original messages, it became an effective tool for shaping the campaign narrative by transmitting posts from other platforms to coordinate political party communication, including content that violated electoral laws.

Of the total number of electoral violations observed, only 4% of Sinhala posts and 6% of Tamil posts could be identified as being first published on WhatsApp. Most of the observed posts had been posted on Facebook and then shared as links or screenshots in WhatsApp groups. Others included YouTube videos, X posts, and a few TikTok posts. This suggests that WhatsApp is used as a vehicle for harmful content, contributing to amplification of harm across platforms.

Figure 5: Origin of posts (WhatsApp versus others)

`

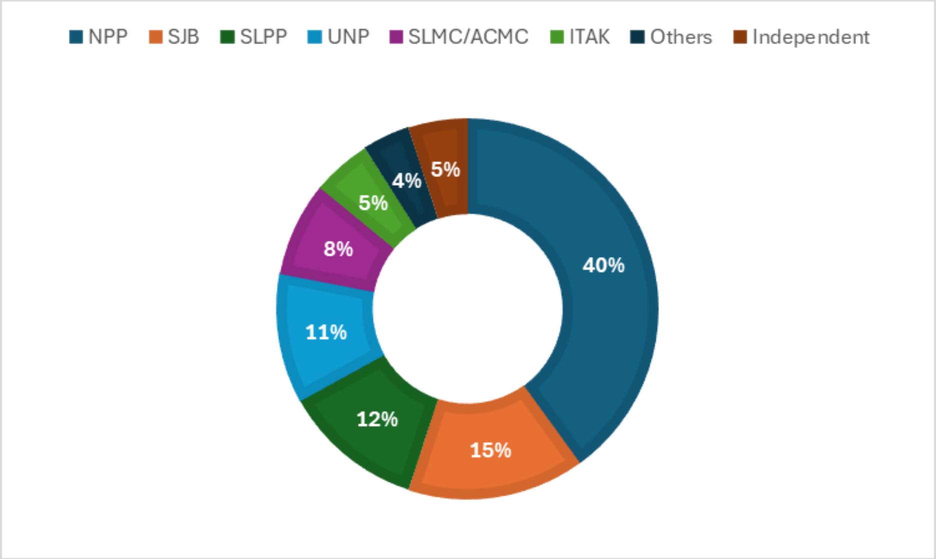


For the purpose of study, the total number of 1753 WhatsApp groups and channels that were considered for data collection, were analysed in terms of their affiliation with a political party or otherwise (such as showing no specific political party-bias).

The analysis showed that the majority of political content shared in WhatsApp groups was aligned with the National People's Power (NPP), which accounted for 40% groups. The main opposition party, Samagi Jana Balawegaya (SJB) followed with 15% groups, while the Sri Lanka Podujana Peramuna (SLPP) and the United National Party (UNP) maintained 12% and 11% groups respectively. Smaller but still notable presences were observed among Sri Lanka Muslim Congress (SLMC) and All Ceylon Makkal Congress (ACMC) affiliated groups, which together accounted for 8% groups, while Illankai Tamil Arasu Kachchi (ITAK) maintained 5 groups. A further 4% groups fell into the "Others" category, reflecting affiliations with minor parties, while 5 %were classified as independent, with no explicit party alignment. Overall, the distribution shows that while multiple parties maintain a WhatsApp presence, the NPP has developed the most extensive and organised digital footprint across content-sharing groups.

Figure 6: Share of observed WhatsApp groups by political party affiliation

`

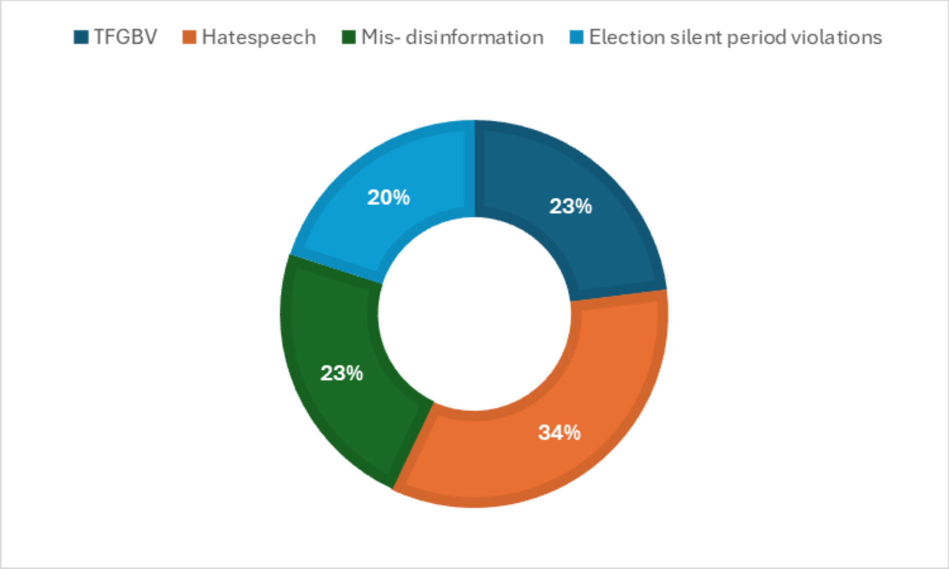


3.3. Thematic Analysis

3.3.1. Major types of electoral violations

A thematic analysis of reported posts reveals a pattern of four types of key electoral violations: Hate speech and defamation (34%), mis- and disinformation (23%), technology facilitated gender-based violence (23%), and violation of election laws during the silent period (mainly silent period violations (20%). These posts were typically not directed at ordinary end users, but at campaign operatives engaged in manipulating public opinion and undermining the democratic processes that ensure free and fair elections.

Figure 7: Share of types of violations



`

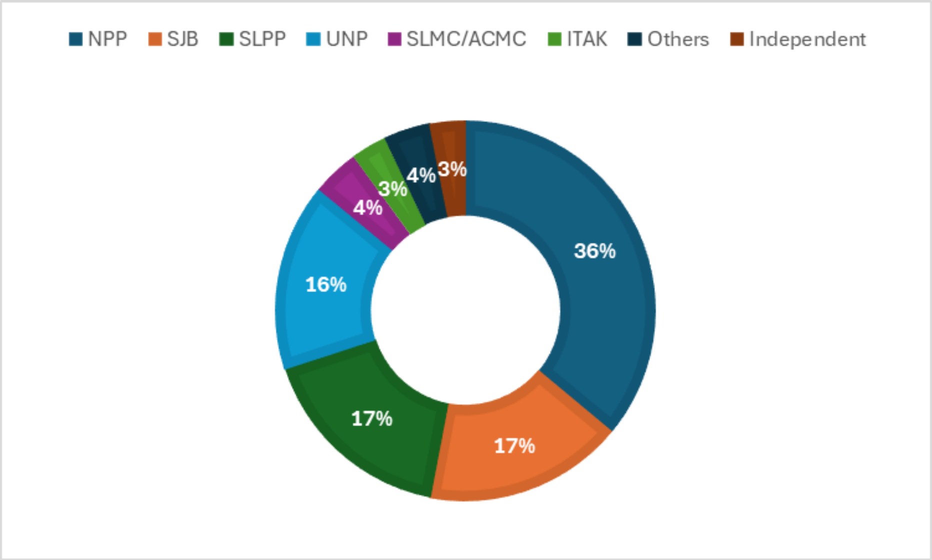
3.3.2. Electoral Law Violations by Groups Affiliated with political parties

The analysis also sought to trace the origin of electoral violations to specific parties or actors by examining the political affiliations of the observed groups and accounts. For WhatsApp groups without a clear party link, violations were assessed based on the content of the offending posts and their previous activity. The following chart illustrates the party or group alignments behind posts that violated election laws.

Reflecting its large-scale, centralized social media mobilization as it surged toward electoral victory, the NPP dominated electoral law violations with 36% of offending posts. In contrast, established parties like SJB, SLPP, and UNP — facing electoral headwinds — each accounted for 17%, 17%, and 16% respectively, suggesting a decline in centralized campaign coordination, with individual candidates and smaller groups largely driving their online presence.

While the limited data does not allow for broad generalization, the comparison between group alignment and electoral violations suggests that group share did not always correspond proportionally to violation share. The NPP, with the largest share of monitored groups (40%), also led in violations (36%), reflecting a broadly proportional pattern — as did SJB, whose 15% group share aligned closely with its 16% violation share. Notably, the majority of NPP-linked violations were concentrated within a small number of WhatsApp groups dedicated to negative campaigning against the opposition, often operating under the guise of 'satire.'

Figure 8: Party and group affiliations of posts violating election laws



Smaller parties such as the SLMC & ACMC (two Muslim dominant political parties) and ITAK showed a nearly identical share of violations, respectively 4% and 3% while they accounted for 8% and 5% of the share of groups. Independent groups also mirrored this pattern. Notably, there was also a small number of violations that came from accounts with vague or unclear affiliations, underscoring the difficulty of tracing the party affiliation of some content. This may trigger personal attacks on identified individuals by self-motivated actors with little association to political parties or groups.

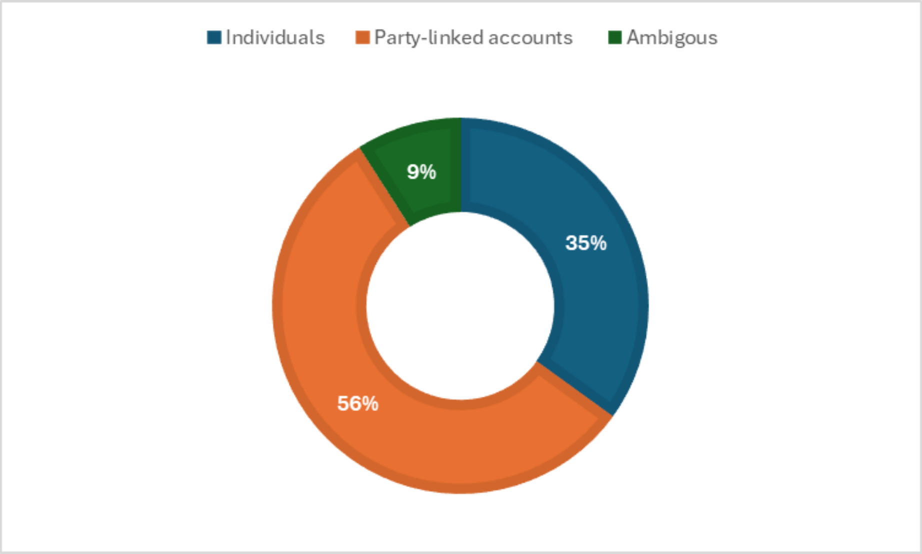
In contrast, groups associated with two political parties displayed a higher share of electoral violations relative to their share of observed groups. For the SLPP, the proportion of violations (17%) exceeded its share of groups (12%), and for the UNP, violations accounted for 16%, indicating a comparatively higher incidence of violations and a greater contribution to TFGBV and hate speech. We do not have data-driven evidence to explain this pattern. However, given the electoral setbacks these two major parties experienced in recent elections, a more aggressive approach to campaigning may have contributed to the higher incidence of violations.

3.3.3. Silent Period Violations

Sri Lanka’s election silent period begins 48 hours before polling day, and breaches of this period accounted for the largest share of unique election law violations. What is particularly concerning is the organised nature of these violations. These largely fell into three categories: posts by individual candidates or their close supporters promoting a candidate; posts that continued party-level campaigning, often spreading defamatory content against rival parties or groups; and posts whose origins were ambiguous, making it difficult to determine whether they were linked to a particular party or to individual actors.

Figure 9: Source of violations during the silent period

`



Notably, over half (56%) of silent period violations originated from accounts linked to general party campaigns rather than individual candidates. These violations included leadership speeches and videos, circulated widely online through official or affiliated accounts. Many posts promoted general party achievements or propaganda, showcasing welfare programs, development work, or economic progress attributed to the party. Some content took the form of defamatory or fearmongering narratives, designed to discredit rival parties or create anxiety about their policies. In addition, parties engaged in centralised messaging, using hashtags, slogans, memes, and coordinated visuals spread across multiple groups and platforms to sustain visibility during the silent period.

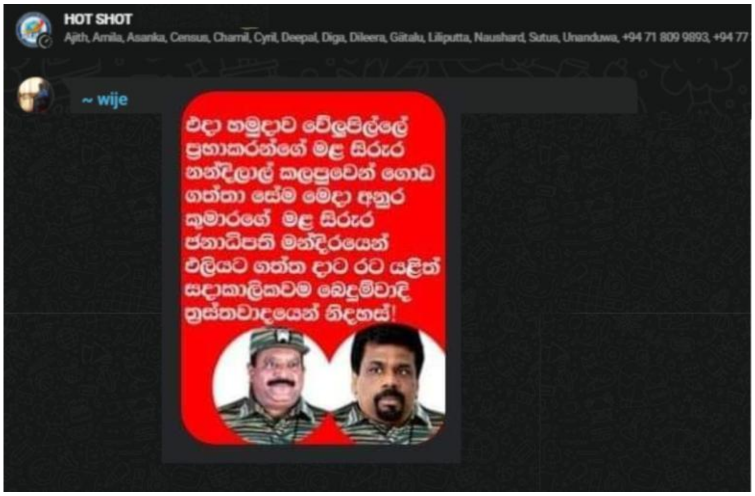
In contrast, posts outside this organised pattern usually promoted individuals and tied to specific localities. Some posts were self-promotional, highlighting personal achievements, promises, or qualities to attract voter support. Others were tied to localised campaigning, such as promoting community events, neighbourhood visits, or small gatherings within a constituency. Candidates and supporters also made targeted appeals to identity groups, framing messages around ethnic, religious, gender, or youth identities to mobilise voters. There was also negative campaigning at the constituency level, where rival candidates were directly attacked, often in personalised and hostile terms.

Some violations could not be clearly traced to candidates or parties. In other cases, special interest groups such as women's organisations or religious networks amplified partisan content, often extending its reach beyond local boundaries. n this category, third-party influencers and media pages also played a role, sharing or producing partisan messages despite not being formally affiliated with political parties, effectively acting as campaign proxies and potentially circumventing silent-period restrictions.

3.3.4. Hate Speech and Defamation

Hate speech in the observed WhatsApp groups operated at two distinct levels. At the communal level, posts framed and vilified ethnic and religious communities, most notably Muslims and Tamils, portraying them as enemies, reinforcing hostile stereotypes, and deepening divisions. At the individual level, political leaders were personally targeted through derogatory language, slurs, mocking images, and insulting nicknames. These attacks went beyond political critique, portraying opponents as illegitimate or contemptible figures and shifting the discourse from policy debate to personal denigration.

Often these narratives are combined to attack personalities. For example, minority voters (Muslim and Tamils) in the Colombo Municipal area were framed as '*sakkiliya'* (a very unhygienic or uncultured person) for supporting the NPP's Mayoral candidate of Colombo Vraie Balthazaar.[17] Hate speech also took the form of leadership-targeted attacks, directed disproportionately at party leaders. Notably, President Anura Kumara Dissanayake, Prime Minister Harini Amarasuriya, and Opposition Leader Sajith Premadasa were singled out for ridicule and vilification, even though some of them were not direct contestants in the local government elections.



HOT SHOT
Ajith, Amila, Asanka, Census, Chamil, Cyril, Deepal, Diga, Dileera, Gàtalu, Liliputta, Naushard, Sutus, Unanduwa, +94 71 809 9893, +94 77

~ wije

එදා හමුදාව වේලුපිල්ලේ
ප්‍රභාකරන්ගේ මළ සිරුර
නන්දිලාල් කලපුවෙන් ගොඩ
ගත්තා සේම මෙදා අනුර
කුමාරගේ මළ සිරුර
ජනාධිපති මන්දිරයෙන්
එළියට ගත්ත දාට රට යළිත්
සදාකාලිකවම බෙදුම්වාදි
ත්‍රස්තවාදයෙන් නිදහස්!

A post circulating on a WhatsApp group calling for the death of President Anura Kumara Dissanayake. Direct translation reads:

"Just as the Army once dragged the body of Velupillai Prabhakaran from the Nandikadal lagoon, the country, too, will be freed from separatism when the body of Anura Kumara is taken out of the Presidential

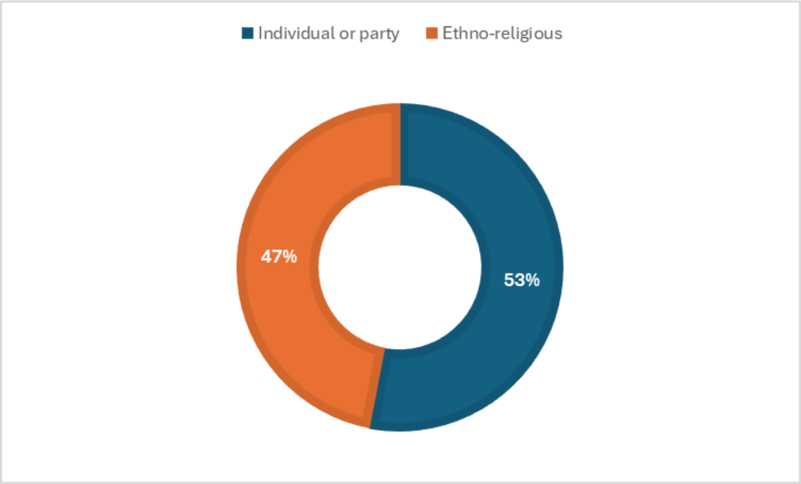
[17] A close translation of the word is accessible on: https://hatebase.org/vocabulary/sakkiliya.

`

In the case of Harini Amarasuriya and Colombo Mayoral candidate Vraie Balthazaar, the discourse included gendered defamation. Attacks against them were not limited to political critique but extended to sexist insults and personal ridicule, highlighting the particular vulnerability of female politicians to gender-based hate speech in online spaces.

Another common feature was enemy framing, where political opponents, LG candidates and party supporters were depicted not merely as rivals but as hostile actors who should be treated with contempt. Such framing reinforced the idea that adversaries were illegitimate, corrupt, or even dangerous, fuelling divisive narratives.

Figure 10: Ethno-religious versus party/individual oriented hate speech

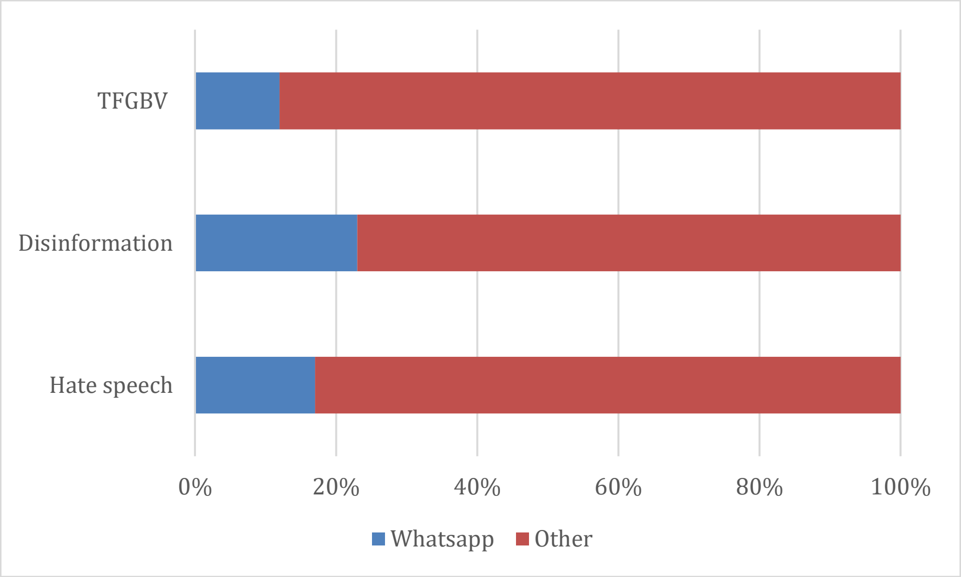


The chart illustrates the two broad categories of hate speech observed during the monitored period. The majority of violations (53%) fell under the "individual or party" category, referring to targeted attacks, defamation, or incitement directed at specific political figures or parties. The remaining 47% constituted ethno-religious hate speech. This includes content that exploits ethnic, religious, or communal identities to provoke division or hostility, that indicates a particularly concerning finding given Sri Lanka's history of ethno-religious conflict. The near-equal split between the two categories signals that hate speech during this electoral period was both politically motivated and communally charged.

Another sub-category was the use of polarising labels. During the campaign, narratives framed candidates of the governing party as “liars” and opposition candidates as “thieves.” These simple but powerful labels entrenched hostility between groups, deepening the axis of political polarisation in online discourse.[18] Some of them may not qualify as 'prohibited/ regulated speech' under the ICCPR convention.

The spread of this discourse was amplified through cross-platform circulation. Content such as Facebook posts and Facebook reels labelling politicians as "liars" or "thieves" often migrated into WhatsApp groups, where they gained further traction. Often Whatsapp groups became the vehicles for Facebook content to be shared as Facebook links. This cross-platform flow enabled hate speech to spread more widely, reinforcing negative narratives across digital spaces.

Figure 11: Origin of violations



Confirming the observation in Sri Lanka and elsewhere, that WhatsApp is often used as a vehicle for materials produced on other platforms, when it comes to electoral violations, data indicates that a significant share of election-related violations were reproduced and circulated through WhatsApp.[19] These often originated from other platforms such as Facebook, TikTok or X. In

[18] Harindra Dassanayake and Rajni Gamage (2025). *Horu versus Boru: The politics of the 2025 local government elections in Sri Lanka*. Polity. Retrieved from https://polity.lk/harindra-b-dassanayake-and-rajni-gamage-horu-versus-boru-the-politics-of-the-2025-local-government-elections-in-sri-lanka/.

[19] A 2024 NYU report documented how platforms like Telegram use bots to automatically share content to X (Twitter), and apps like India's ShareChat enable users to cross-post content directly onto WhatsApp and other platforms — creating what researchers describe as "feedback loops" where the same content keeps "popping up in different parts of the platform ecosystem" (Euronews/NYU, 2024). Full source: https://www.euronews.com/next/2024/10/10/report-shows-how-messaging-apps-are-used-to-spread-political-propaganda.

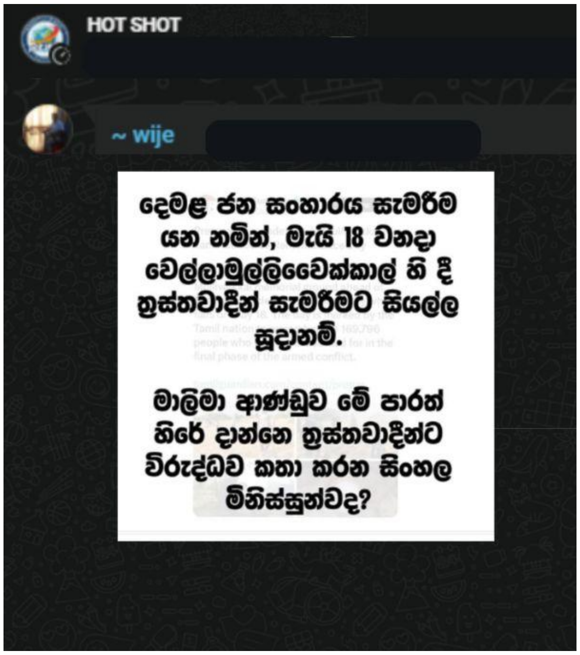
Further, a peer-reviewed study analyzing the 2022 Brazilian elections across WhatsApp, Twitter, and Kwai found that misinformation content routinely migrated across platforms, being "reshaped or recontextualized to fit the new platform's affordances", confirming WhatsApp's role as a secondary redistribution channel for content originating

`

cases where the original source platform could not be clearly identified, the violation was attributed to WhatsApp. This highlights WhatsApp’s role as both a direct site of violation and a key channel for amplifying harmful content from other online spaces.

3.3.5. Disinformation and Fake News

Local government elections campaigns and the period immediately following the election saw a considerable amount of fake news. Often the content took the form of attacks, aiming to discredit or harm a political opponent, or propaganda, seeking to bolster a party's image. A major theme was the deliberate spread of false or misleading information. In certain instances, these were used to push narratives aimed at discrediting a political figure, deliberately twisting or misquoting a statement that they had made. Some other posts that were shared in WhatsApp groups during the LG election time showed unrelated pictures of politicians along with negative news, such as criminal activities.



HOT SHOT [redacted]

~ wije [redacted]

දෙමළ ජන සංහාරය සැමරීම
යන නමින්, මැයි 18 වනදා
වෙල්ලාමුල්ලිවෙක්කාල් හි දී
ත්‍රස්තවාදීන් සැමරීමට සියල්ල
සූදානම්.

මාලිමා ආණ්ඩුව මේ පාරත්
හිරේ දාන්නෙ ත්‍රස්තවාදීන්ට
විරුද්ධව කතා කරන සිංහල
මිනිස්සුන්වද?

This post mislabels theTamil communities commemorating the Mullivaikkal massacre (May 18) — a recognized day of mourning for Tamil civilians killed in the final stages of Sri Lanka's civil war — as "extremists," and uses inflammatory language to provoke Sinhalese nationalist sentiment. The closing question is designed as a call to action, framing ethnic remembrance as a threat. The translation reads:

"To commemorate the Tamil genocide, from May 18th, all Mullivaikkal extremists are ready to commemorate.
Is it the Sinhalese people who speak against these extremists who are going to be arrested by this government again?"

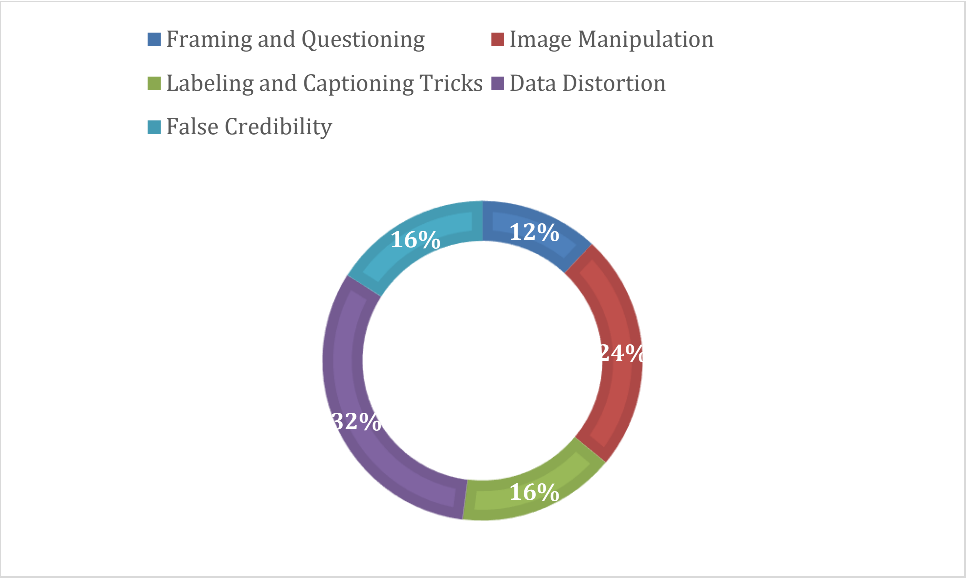
Some of the techniques used in WhatsApp-based sharing of fake news and misinformation include raising rhetorical questions about political figures, posting mismatching or decontextualized pictures, and deliberately sharing a fake news item with the label “fake” to create doubt in the minds of users. Other tactics observed were the use of satirical or mocking

elsewhere (*International Journal of Public Opinion Research*, 2024). Full source: https://academic.oup.com/ijpor/article/36/3/edae032/7709027 captions to disguise disinformation as humour, recycling old content as if it were recent, exaggerating or fabricating statistics, and forwarding unverifiable claims attributed to anonymous “insiders” or “sources” to lend false credibility. These strategies were designed to confuse users, exploit existing biases, and amplify distrust in political processes. An analysis of the generation of fake news and misinformation on WhatApp groups identified the following typology:

a. Framing and Questioning (rhetorical framing, satirical presentation)
b. Image Manipulation (mismatching images, recycled content)
c. Labeling and Captioning Tricks (false labeling, fabricated captions)
d. Data Distortion (fabricated or exaggerated statistics)
e. False Credibility (anonymous or unverifiable sources)

A breakdown of these various types of violations is illustrated below.

Figure 12: Types of misinformation in observed WhatsApp groups

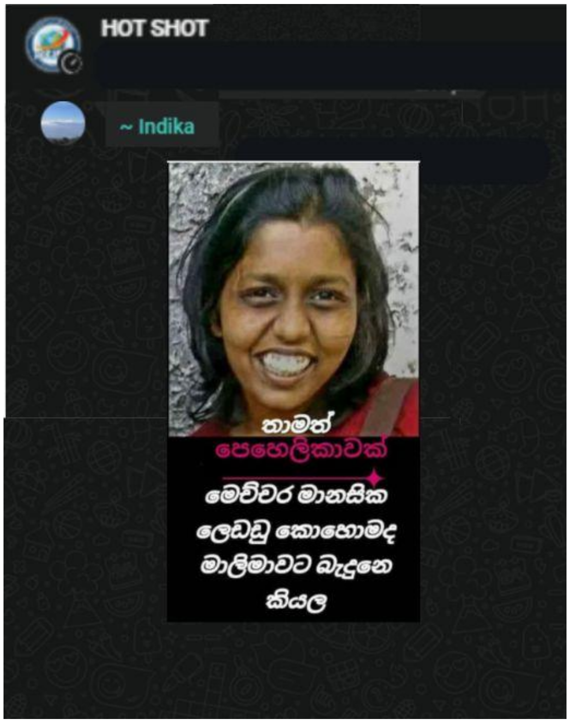


Based on observable patterns such as timing, similarity of messaging, and cross-platform amplification, some of the content appears to have been deliberate and potentially coordinated. As discussed in the next section, these posts and related propaganda circulated in ways that suggest networked dissemination across different parts of the country.

Technology-Facilitated Gender-Based Violence (TFGBV)

`

Technology-Facilitated Gender-Based Violence formed one of the key challenges on social media, and WhatApp groups were no exception. During Sri Lanka’s 2025 Local Government elections, gender-based and anti-LGBTQI+ disinformation and harassment were widespread. Prime Minister Harini Amarasuriya, though not a candidate, became a frequent target of misogynistic, sexist, and homophobic attacks. A viral video and smear campaign falsely linked her and her coordinating secretary to an alleged same-sex relationship, while other posts weaponized the NPP’s progressive rhetoric on sex work through gendered slurs.



HOT SHOT

[redacted]

~ Indika

[redacted]

තාමත්
පෙහෙලිකාවක්
මෙච්චර මානසික
ලෙඩඩු කොහොමද
මාලිමාවට බැදුනෙ
කියල

Using the photo of MP Lakmali Hemachandra, the post uses dehumanizing language to publicly mock and discredit her. It targets her appearance and mental fitness to shame her, a tactic commonly weaponized against women in public life. The translation reads:

Still a puzzle: How can such mental patients joined the NPP.

One feature of TFGBV is that these allegations and campaigns show a long shelf-life. Posts from previous election campaigns often resurfaced. A non-consensual leak of intimate images of a female NPP supporter, along with videos harassing Amarasuriya and activist Deepani Silva, directly violated the Election Commission’s Code of Conduct.

Even with fewer women contesting, derogatory terms were used against female candidates, alongside continued targeting of LGBTQI+ individuals. While NPP fielded more female candidates, they were the primary targets of gender-based attacks, similar tactics were also directed at UNP, SJB, and SLPP LG candidates. These narratives framed women councillors as sexual objects and spread claims that they offered sexual bribes, reinforcing misogynistic stereotypes and undermining the political legitimacy of these women. In the Northern and Eastern regions, TFGBV was particularly directed at Tamil-speaking communities, while

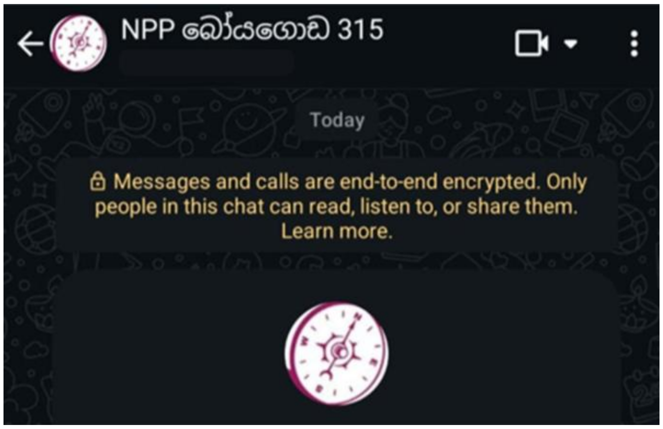
Muslim candidates were mocked online with claims that their identities could not be verified due to face coverings.

3.4. WhatsApp Enables Coordinated Messaging Campaigns

3.4.1. Cascading WhatsApp Groups

WhatsApp facilitates coordinated campaigns through what can be described as a cascading system of groups. By this, we refer to an arrangement in which multiple WhatsApp groups are linked indirectly through shared administrators or members, allowing content posted in one group to be systematically forwarded to others in a step-by-step sequence. In several instances observed at the local level, groups were named systematically or numbered, suggesting their placement within a broader communication structure. Each group functioned as a distribution node: messages, instructions, or campaign material introduced in one group would be relayed to parallel or subordinate groups, thereby expanding reach incrementally. This cascading flow—moving from a central or core group to successive layers of groups—enabled rapid and cost-effective dissemination of information (or disinformation), while preserving a degree of coordination and oversight over how messages travelled across networks. The resulting structure resembles a tiered, networked communication system that can be scaled quickly for organised outreach or mobilisation.

Figure 13: WhatsApp Groups mirroring public administration system's coding



NPP බෝයගොඩ 315

Today

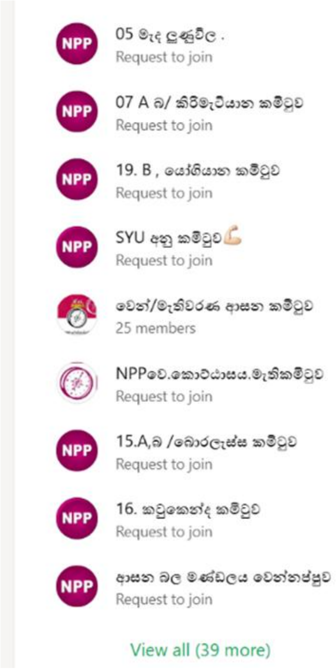
Messages and calls are end-to-end encrypted. Only people in this chat can read, listen to, or share them. Learn more.

These groups were often organized by village or Grama Niladhari division and frequently named similarly to the official numbering of the Grama Niladhari division, conflating political party and state structures. In the post above, the *Boyagoda* WhatsApp Group was organised alongside the

`

registration number of the Grama Niladhari division in the Ministry of Public Administration Provincial Councils and Local Government of Sri Lanka. This can also confuse citizens, as the naming and organization mirror official administrative structures, creating the impression that political messaging is state-sanctioned.

Figure 14: Organisation of WhatsApp groups in a cascade system



05 මැද ලුණුවිල .
Request to join

07 A බ/ කිරිමැටියාන කමිටුව
Request to join

19. B , යෝගියාන කමිටුව
Request to join

SYU අනු කමිටුව💪
Request to join

වෙන්/මැතිවරණ ආසන කමිටුව
25 members

NPPවෙ.කොට්ඨාසය.මැතිකමිටුව
Request to join

15.A,බ /බොරලැස්ස කමිටුව
Request to join

16. කටුකෙන්ද කමිටුව
Request to join

ආසන බල මණ්ඩලය වෙන්නප්පුව
Request to join

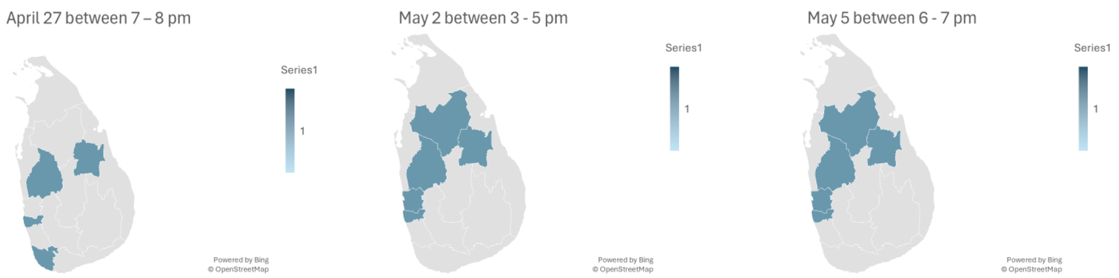
The NPP mobilized this system effectively, using the networked WhatsApp groups to coordinate messages and outreach, while other parties lacked a similarly organized and systematic approach.

3.4.2. Simultaneous Messaging

During the election period, WhatsApp was effectively used to distribute large amounts of political content simultaneously. A few fake news items appeared simultaneously across multiple districts in the country, within observed WhatsApp groups. The highlighted districts on each map, corresponding to different dates and time windows, show a coordinated pattern of dissemination, indicating the presence of a networked, simultaneous messaging system. This cascading network allows content to spread quickly across geographically dispersed regions, reflecting how information, or disinformation, can be efficiently amplified through WhatsApp.

Figure 15: Simultaneous messaging on WhatsApp groups

`



While this data is insufficient to establish causality, the patterns provide important preliminary insights into the mechanics of coordinated messaging and the potential reach of these networks across multiple districts.

This study shows that grassroots-level electoral violation monitors play a crucial role precisely because of their embeddedness in local communication networks. They can intercept and document violations that circulate within closed WhatsApp groups and other private channels that are invisible to national observers or platform-level monitoring tools. Their physical and social proximity to these localities gives them access to information that never surfaces in the public domain, making their on-the-ground presence an indispensable layer of accountability and electoral integrity.

4. Conclusions and Recommendations

4.1. Key Observations

WhatsApp functioned as a highly effective, though discreet, tool for political campaigns in Sri Lanka during the Local Government Elections in 2025. Unlike public platforms designed for broad visibility, its strength lies in enabling precise, coordinated messaging across multiple regions, often structured in a hierarchical, cascade model. In this setup, campaign organizers or influential community figures distribute messages to a small number of trusted "sub-leaders", who then forward content to their local groups, and so on. This creates a rapid amplification chain, where information, whether accurate or deliberately misleading, spreads quickly through trusted social networks, making recipients far more likely to believe and share it.

The platform's end-to-end encryption, combined with the large volume of activity in private groups, makes these exchanges largely invisible to external observers, researchers, and regulators. Unlike open social media channels, there is minimal moderation or public accountability, meaning that fake news, inflammatory material, and targeted disinformation can circulate largely unchecked. The result is the creation of segmented political echo chambers, where manipulative narratives are reinforced and dissenting voices are rarely present, giving campaigns a highly effective, low-cost, and low-risk channel for influencing opinion.

4.2. Recommendations

1. Election regulators should actively engage with WhatsApp groups, channels, and communities to enhance transparency and accountability.
2. The platform's leveraging capacity, including the creation of large groups and communities, should be assessed and contextualized within WhatsApp's terms of service and regulatory framework.
3. Awareness-building initiatives should target political actors, civil society, and election monitors to highlight the potential misuse of WhatsApp in political processes.
4. Comprehensive election monitoring on WhatsApp should be implemented to track disinformation, abusive content, and coordinated messaging campaigns.

`

MURAGALA | CPPP
CENTRE FOR PROGRESSIVE POLITICS & POLICY

Published in February 2026

About Muragala | Centre for Progressive Politics & Policy

Muragala | Centre for Progressive Politics & Policy (CPPP) is a politics and policy-oriented research collective established in 2023, which promotes equal & equitable societies in Sri Lanka & the region. Our work lies at the intersection of political science, political economy, and political sociology. Together, we aim to generate ideas, enrich the discourse, and mobilise social action to create a more robust Global South discourse and politics.

Acknowledgments

This report was written on behalf of Muragala | CPPP by Harindra B. Dassanayake, Nipunika Pathirana, Sudaraka Arthanayake, Sajeewa Wijeweera, and Manuja Wijesuriya, based on data collected and reported by 30 grassroots activists from different districts across the country, as part of the *Strengthening Resilience Against Disinformation in Sri Lanka* project. This publication was co-funded by the European Union. Its contents are the sole responsibility of the authors and do not necessarily reflect the views of the European Union.

Co-funded by
the European Union

DEMOCRACY
REPORTING
INTERNATIONAL

factum