India: The dissemination of misinformation on WhatsApp is driving vigilante violence against minorities

Shakuntala Banaji and Ram Bhat

Over the last few years, hundreds of people have been killed or injured by vigilante mobs across India. The rapid spread of misinformation (ranging from unintentional deception to deliberate disinformation) via the use of digital media applications such as WhatsApp has played a crucial role in enabling the rise of this violence, including lynchings.

While these incidents have included random strangers being beaten to death on the suspicion that they are potential child-kidnappers or organ-snatchers, discriminated groups such as Muslims and Dalits have been especially targeted, particularly in relation to allegations of cow slaughter – an issue that has increasingly been used as a catalyst for attacks by right-wing Hindu nationalists.

Typology of misinformation

In order to better understand how WhatsApp and other social media messaging platforms are implicated in discriminatory mob violence, using funding from WhatsApp we conducted an independent qualitative study in four large states of India between November 2018 and August 2019. As part of this study, we interacted with nearly 300 WhatsApp users from a wide range of backgrounds: men and women aged between 18 and 50, in both rural and urban areas, from upper and lower castes, and including Hindus, Muslims and Adivasis with a variety of political beliefs and occupations. We also studied more than 1,000 WhatsApp anonymized messages that were typically shared in WhatsApp groups.

A man holds a placard that reads 'stop attacking minorities' during a protest against mob lynching. Kolkata, India.

Credit: SOPA Images Limited / Alamy

Based on this review, we developed a typology of violence-fuelling misinformation that is most commonly received and shared by Indian WhatsApp users. The categories include:

• **Overwhelming content:** still and moving images of incidents from across the world, shared without context, each displaying something spectacular – an execution, an accident, a child getting beaten up, a natural disaster, fires and so on – to engage users by imparting a sense of shock. This content also serves the function of establishing a WhatsApp group as a significant channel of information unavailable in mainstream broadcast and print media, where graphic violence is generally not shown.

• **Nationalism and ethno-religious bigotry:** these messages are conspiratorial, drawing on a wide number of established stereotypes and prejudices against minority populations. They often build on negative propaganda featured in mainstream outlets, such as anxieties around population growth among poor and minority populations, conspiracy theories about the forced conversion of disenfranchised Hindus to Christianity, smear campaigns aimed at opposition politicians, and other narratives aiming to incite hostility towards particular groups or individuals.

• **Miscellaneous:** this includes festival-related greetings, videos of animals, television clips from talent shows and news programmes, public events, humorous clips from India and across the world (including content imported from other platforms such as YouTube and TikTok) and other material. Though seemingly innocuous, these snippets function to sustain the impression of a constant 'flow' of information and to build the profile, brand and

legitimacy of users who pass on other kinds of misinformation.

- *Gendered content:* as in many other areas, WhatsApp usage in India continues to be highly gendered. Women across age, religion, class and caste do not have unrestricted access to mobile phones, and their use is often closely monitored and controlled by their husbands, brothers, fathers or other male relatives. Young women who do use social media frequently experience messages threatening them with devastating physical and symbolic violence, including rape, harassment and the public sharing of personal information to intimidate them into silence. Frequently, incidents of online harassment result in a vicious circle, where families and communities blame the women for attracting these attacks.

A recurrent trope that has recently emerged is reports — usually a short video or an image with a voiceover — of child-kidnapping 'gangs' or an individual kidnapper allegedly doing the rounds in or around a community. Since 2017, when the frequency of these attacks dramatically increased, dozens of people have been lynched on the suspicion that they are child-kidnappers. Usually, in South Indian states the 'stranger' is described as being from North India and vice versa. More recently, the same misinformation is being shared with the putative strangers now being described as Rohingya Muslim refugees, thus playing into the government's strategic generation of fear and loathing of outsiders in the wake of its National Register of Citizens initiative.

User motivations

A significant section of WhatsApp users expressed fatigue with the volume of WhatsApp messages received in a single day. They were members of several groups formed on the basis of family, friendship, neighbourhood, caste, religion, political party and occupation. Users were part of most of these networks offline and active online participation cemented their credibility and membership in the offline world. It mattered to users that they be seen as active, knowing WhatsApp participants in specific groups.

Some users — particularly those over 25 — categorized messages on the basis of sender-credibility. Younger users, on the other hand, were more sceptical about the accuracy of the messages because of enhanced functional media literacy. If a video message was edited heavily, for instance, they were able to spot the places where the video had been altered and were suspicious of its authenticity — although this did not always lead them to reject the message or to report it. Most users suspended their scepticism during politically charged moments such as cross-border conflict and general elections, or regarding news of child-kidnappings. During such occasions, users reported that they forwarded messages out of an assumed sense of civic duty, and out of loyalty to family or communal ties that have historic roots. That these ties were often caste-based, partisan and led to the spread of misinformation was less important to them than displaying adequate nationalist or communal fervour at a fraught moment. Another finding was that, contrary to popular belief, users with little

The cover of Indian magazine 'India Today' with the headline 'The Weaponisation of Whatsapp'.

Credit: Nick Kaiser/dpa/ Alamy

or no digital and media literacy played a minimal role in the spread of misinformation. A very small number would receive and forward misinformation uncritically. However, it was primarily upper-caste privileged men, and some women connected to them, with high levels of functional media literacy and class capital, who were involved both in producing this content and creating the networks to disseminate it to others.

Inadequate state response

The increasing use of media platforms such as WhatsApp to harass citizens who dissent from the ruling party agenda and its nationalistic ideology has been enabled by impunity. Not only have the police and security forces failed to prevent the spread of misinformation but there is also evidence that a significant portion

may themselves sympathize with the prejudices the messages express. In one 2019 survey of 12,000 police personnel conducted across 21 Indian states, a third of respondents felt that mob violence by Hindu vigilante groups against Muslims in cases of alleged cow slaughter was 'natural', while more than half felt that complaints of gender-based violence were false.

Given the institutional failure of the state and law enforcement agencies to control violence related to misinformation, the central government has proposed that platforms such as WhatsApp allow users to become traceable, in order to identify users who share misinformation so action be taken against them. In effect, the government has proposed that encryption be removed from WhatsApp in the name of law and order: the case is currently being heard by the Supreme Court of India. However, given that far-right and nationalist groups continue to disseminate misinformation with impunity on public platforms such as Facebook, ending encryption is unlikely in itself to resolve the problem of online incitement.

Furthermore, human rights activists have expressed concern that the government could exploit such a move to monitor and repress political opponents and dissenters, rather than curb hate speech against minorities and other marginalized groups. For instance, the district-level administration in different parts of the country has arbitrarily placed restrictions on WhatsApp usage during sensitive periods. For example, in the districts of Kargil and Leh, internet access was restored in December 2019

after a five-month internet shutdown in the erstwhile state of Jammu and Kashmir. With the restoration of internet services, the district administration has demanded that all administrators of WhatsApp groups are required to register themselves with the district administration, with strict orders that they will be held responsible for any content circulated on these WhatsApp groups.

This has taken place against the backdrop of the Bharatiya Janata Party (BJP) government's growing authoritarianism, particularly towards Muslims and democratic dissenters. The erstwhile state of Jammu and Kashmir and the state of Uttar Pradesh, both states with large Muslim populations, have faced large-scale violence and the imposition of curfews, internet shutdowns and police aggression against protesters. In late 2019, the central government passed the Citizenship Amendment Act (which came into operation on 10 January 2020), providing citizenship to persecuted minorities of different religions from neighbouring countries except Muslims. In addition, the central government has also allocated budgets for undertaking the National Population Registry that seeks to identify 'valid' citizens on the basis of identity documents. There seems to be little doubt that the Citizenship Amendment Act, in conjunction with the National Population Register and National Register of Citizens, could be used to label vulnerable Muslims as foreigners or illegal 'infiltrators'. Facing popular protest against the Citizenship Amendment Act throughout the country, the BJP and the larger family of Hindutva organizations have resorted

to using WhatsApp to incite anti-Muslim hatred among its supporters.

Even as the central government has weaponized its administrative and military powers against Muslims, especially in Kashmir, the same officials have claimed that encryption prevents them from acting against those spreading misinformation. It is difficult to believe this argument is made in good faith, given the lamentable track record of the government and security agencies to tackle hate speech on public social media platforms such as Facebook, Twitter, Instagram and YouTube. In spite of repeated complaints by Dalit activists, feminists, journalists, human rights activists, students, academics and many others, no action has been taken by those who perpetrate death threats, sexual harassment and other forms of abuse. Given that both technology companies and law enforcement agencies have failed to act against hate speech on open platforms, it is doubtful that the removal of encryption will result in any action.

Towards real change

Vigilante violence is linked to specific typologies of misinformation produced, shared by and targeted against specific social groups, along the lines of caste, indigeneity, class, religion, gender, sexuality, language and ethnicity. This violent social structure should be taken as the broad context in which applications such as WhatsApp are used. WhatsApp usage further intensifies this violence in specific mediated ways. Motivations of users are important since users have diverse ways to

A smartphone shows no network available under the Kashmir communications blackout.
Credit: Saqib Majeed/ZUMA Press

justify misinformation that range from suspension of disbelief to civic duty and nationalism. Such ideological articulations are or should be contested in order to transform the conditions under which social interactions are imagined and acted upon.

Unfortunately, mainstream thinking on curbing misinformation has been restricted to purely technical solutions, such as possibly removing encryption and investment in functional media literacy. However, misinformation and subsequent vigilante violence need to be curbed through a critical understanding of broader socio-political contexts. Interventions such as promoting a stronger media literacy around the politics of representation and power, as well as cross-stakeholder cooperation to act on hate speech, would be important steps towards real change.

More details of the research project and its findings are available in S. Banaji and R. Bhat, WhatsApp Vigilantes: An Exploration of Citizen Reception and Circulation of WhatsApp Misinformation Linked to Mob Violence in India, LSE, 2019.