

The growth of online hate speech in Sri Lanka does not guarantee another pogrom. It does however pose a range of other challenges to government and governance around social, ethnic, cultural and religious co-existence, diversity and, ultimately, to the very core of debates around how we see and organise ourselves post-war.

Liking violence

A study of hate speech on
Facebook in Sri Lanka

Shilpa Samaratunge and Sanjana Hattotuwa
Centre for Policy Alternatives | September 2014

Contents

Introduction.....	3
Executive Summary.....	7
Overview	9
Methodology.....	11
Scope of Study.....	12
Significance of study	14
Hate speech versus free speech: Some pertinent frameworks.....	17
Definitions	20
Hate speech & Dangerous Speech	21
Forms of online hate	21
Targets and producers of hate speech.....	24
Demographic of producers.....	26
Main targets of hate speech.....	28
Secondary targets	32
Five defining variables.....	39
The speaker.....	39
The audience.....	40
Content of the speech	43
Socio-historical context	43
Mode of transmission	46
Common characteristics of hate speech used to garner support.....	49
Education	49
Denialism	50
Building Group Solidarity	51
Hero Narrative.....	53
Nationalism	54
Scare Tactics	55
Othering.....	56
Addressing online hate speech in other countries and regions	58
Recommendations	61
References	65
Annexure.....	67
Facebook Page 1.....	67
Facebook Page 2.....	84
Facebook Page 3.....	102

Facebook Page 4	111
Facebook Page 5	120
Facebook Page 6	131
Facebook Page 7	140
Facebook Page 8	151
Facebook Page 9	157
Facebook Page 10	161
Facebook Page 11	162
Facebook Page 12	168
Facebook Page 13	170
Facebook Page 14	182
Facebook Page 15	184
Facebook Page 15	192
Facebook Page 17	194
Facebook Page 18	198
Facebook Page 19	200
Facebook Page 20	202

Introduction

“Hate speech” on the Internet is a global concern and with no kill-switch solution. Depending on the location online, language and media used, context and sometimes even the nature of the actors concerned, dealing with hate speech is a vexed challenge from parent to policymaker. This hasn’t stopped politicians, with little to no understanding of underlying technical challenges or repressive governments, who often seek a monopoly around the dissemination of defamatory propaganda seeking to control hate speech. Parochialism and expediency drive most efforts around hate speech related policy responses and legislation. In Sri Lanka, online social media and web based platforms, accessed increasingly over smartphones and tablets, provide an important, necessary vent for critical dissent, in a context where mainstream media does not and cannot afford the space for questioning or content that holds the government accountable for heinous crimes and outrageous corruption. The growth of content creation and consumption online, wider and deeper than any other media in the country and at an accelerated pace, has also resulted in low risk, low cost and high impact online spaces to spread hate, harm and hurt against specific communities, individuals or ideas. Conspiracy theorists, fringe lunatics and trolls have since the first days of the Internet inhabited online spaces and engaged with devoted followers, or sought to deny and decry those who question them. The growth of hate speech can be seen as a natural progression outward from these pockets of relative isolation, and is also pegged to the economics of broadband internet access and the double digit growth of smartphones – an underlying, coast to coast network infrastructure capable of rich media content production and interactive, real time engagement. This infrastructure has erased traditional geographies – hate and harm against a particular religion, identity group or community in one part of the world or country, can for example within seconds, translate into violent emulation or strident opposition in another part, communicated via online social media and mediated through platforms like Twitter, Facebook and also through instant messaging apps for mobiles like iMessage and WhatsApp, in addition to the older SMS technology.

A central challenge around addressing hate speech is that it is technically impossible – given the volume, variety and velocity of content production on the Internet today¹ – to robustly assess and curtail, in as close to real time as possible, inflammatory, dangerous or hateful content just in English, leave aside other languages like Sinhala or Tamil. Once content is produced for the web and originally for a single platform, given user interactions and responses, it often replicates and mutates into other content over dozens of other websites and platforms, making it impossible to complete erase a record of its existence even if the original was taken down, deleted or redacted. This makes it extremely hard to address the harm arising out of hate speech, since there is so much of it around in digital form over so many media.

Another challenge is in defining hate speech. Overbroad legislation risks the law being used to curtail and stifle dissent. Loosely defined laws allow perpetrators of hate speech to get away by referencing the freedom of expression. Policymakers who have to respond to angry communities and individuals who are

¹ <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>

the targets of hate speech, if they are important constituencies, often respond with promises to address a problem they in fact cannot. Internet Service Providers and large corporations like Google, Facebook and Twitter have developed robust guidelines around the content they will allow on their platforms, but these seem to only work best around output that is in English. For example, this brief study is testimony to the sheer volume of hate freely disseminated in Sinhalese on Facebook, even though the company has clear guidelines around such content which includes the banning and blocking of users.

Reflecting the lack of any universal definition of hate speech – content acceptable to or posted lawfully in one country or region can be deemed hateful and unlawful in others, even on the same platform or site - the term is, unsurprisingly, variously defined across leading web companies. Google's YouTube defines it as²,

...content that promotes hatred against members of a protected group. For instance, racist or sexist content may be considered hate speech. Sometimes there is a fine line between what is and what is not considered hate speech. For instance, it is generally okay to criticise a nation, but not okay to make insulting generalisations about people of a particular nationality.

Facebook defines hate speech as³,

Content that attacks people based on their actual or perceived race, ethnicity, national origin, religion, sex, gender, sexual orientation, disability or disease is not allowed. We do, however, allow clear attempts at humour or satire that might otherwise be considered a possible threat or attack. This includes content that many people may find to be in bad taste (ex: jokes, stand-up comedy, popular song lyrics, etc.).

Addressing that it hadn't done enough in the past to address hate speech⁴, Twitter's current rules and policies note that⁵,

Users may not make direct, specific threats of violence against others, including threats against a person or group on the basis of race, ethnicity, national origin, religion, sexual orientation, gender, gender identity, age, or disability.

Add to these varying definitions is that the challenge of defining hate speech – in contradistinction to for example content that is just mildly offensive, distasteful, satirical or acerbic – is deeply rooted in context and expression. What could be a generally accepted turn of phrase used in colloquial speech, when translated into English and out of context, can be seen as hate speech in line with the guidelines noted above. Similarly, hate speech can be easily disguised by resorting to non-English and non-textual expression – or a combination of both. This study has a number of examples where inflammatory and defamatory content against a specific community finds expression and openly resides on Facebook because

² https://www.youtube.com/t/community_guidelines

³ <https://www.facebook.com/help/135402139904490/>

⁴ <https://blog.twitter.com/en-gb/2013/our-commitment>

⁵ <https://support.twitter.com/articles/20169997>

it is in Sinhala, a language that clearly lies outside existing language competencies of Facebook's automated and human curated monitoring frameworks. This brings us to a key challenge around hate speech – it always requires context to understand and address, and increasingly, the intermediaries in both supporting and curtailing the spread of it are corporate entities, not governments. Machine level and algorithmic frameworks to identify and block hateful and harmful content often fail, simply because they flag too many false positives (content erroneously flagged as hate speech) or allow so much of hate speech to pass through (in, as noted earlier, languages other than English) that their core purpose rendered irrelevant. This puts the burden of addressing this content on users themselves, who through reporting mechanisms baked into all the major only social media platforms, can choose to report hate speech with relevant context. Only as effective as the numbers who report hate speech, these reporting mechanisms also take some time to kick-in from the time of submission to the actual deletion or blocking of the original content, page, account or user. At a time of heightened violence, this time lag is unhelpful. There is also no guarantee the (corporate) owner of an app, service, platform or website agrees with the reporting of hateful content. Studies show, for example, significant variance in dealing with hate speech even within Facebook⁶.

This study aims to focus these challenges around the significant growth of hate speech in post-war Sri Lanka, primarily directly against the Muslim community and Islam. The rise of Islamophobia in Sri Lanka is well documented⁷ and shows no signs of abating. Studies on this score are often anchored to the statements by extremist groups in public rallies, and physical acts of violence and intimidation. Equally remarkable though less studied is the growth of hate speech in online social media. As I noted in 2013 after the study of just around four of the most active extremist groups on Facebook⁸,

- Even the most offensive anti-Muslim sentiments and statements have a growing audience and following in web based social media
- That such content has a greater chance of going viral, and influencing real world action, when published in online fora as opposed to mainstream and traditional media
- Content is largely visual in nature, appealing to a demographic as young as 18 (who are still in school)
- Anti-Muslim hate speech is generally, qualitatively more vicious and venomous than anti-LTTE sentiments even at the height of war
- Numbers of those joining these groups is on the rise, and the government is either unaware or unable to address this through counter-narratives and content in support of liberal values, tolerance and religious cohesion.

The focus of this study is to expand on these points. Sadly, the content for the research is growing in abundance. When juxtaposed with the increasing violence against sexual, ethnic and religious minorities and the open celebration of hate speech by groups like the BBS with total impunity, content online risks fanning even greater violence in the future. Even if to date there is little evidence of content online leading to actual physical violence, what is particularly disturbing – given the tens of thousands who are actively

⁶ <http://ohpi.org.au/if-you-cant-recognize-hate-speech-the-sunlight-cant-penetrate/>

⁷ <http://newint.org/blog/2013/04/15/islamaphobia-in-sri-lanka/>

⁸ <http://sanjanah.wordpress.com/2013/02/01/anti-muslim-hate-online-in-post-war-sri-lanka/>

producing and engaging with hate speech – is the radicalisation of youth, as young as 18, to an alarming degree. Though discussions and content are respectively conducted and produced in public fora over platforms like Facebook, the scale and degree of this radicalisation remains ironically hidden to politicians, policymakers and even most parents because of a digital media literacy gap. As I warned in 2013⁹,

Given that the extremists are web savvy, and escape the usual checks on the spread of racist content by virtue of publishing material in Sinhala, it is to be expected that unless serious, meaningful and urgent measures are taken by government, hate will overcome more moderate voices online, and risk spilling over to real world violence on the lines of Black July 1983, against Muslims.

Obviously, the growth of hate speech online in Sri Lanka does not guarantee another pogrom. It does however pose a range of other challenges to government and governance around social, ethnic, cultural and religious co-existence, diversity and, ultimately, to the very core of debates around how we see and organise ourselves post-war. What this study lacks, by design, is a list of solutions to counter the growth of online hate speech. There is simply no panacea, no easy fix or solution in the short term that will effectively curtail the emergence of hate speech online in the future. Indeed, a government that protects instigators of hate is not one that can drive progressive policies around addressing a growing trend this same hate expressed online. Politicians who are digitally illiterate are equally ill-placed to bring about legislation that addresses hate speech even though it may appear to be expedient to do so in light of increasing violence. What this study aims to provide is evidence around what remains an under-appreciated driver of conflict and violence post-war. To acknowledge the scale and depth of the problem is a step beyond an ignorance that it even exists. Moving forward requires all levels of government, private corporations outside of Sri Lanka that host social media content, civil society within the country including the legal community, conscientious individuals and institutions in the diaspora and local ISPs, out of a duty of care for their customer base working in concert to address this explosion of hate speech online. Though it is unclear when and if a concerted, collective approach or a wider study around hate speech in Sri Lankan online fora will be undertaken, this report provides a starting point for informed discussions around how urgently this disturbing phenomenon needs to be studied and remedial measures, to the extent possible, taken.

My sincere thanks and appreciation to Shilpa Samaratunge, the lead author and researcher of this study. Despite being profoundly distressed by what she encountered, Shilpa's sharp eye and intelligence was simply invaluable in matching existing research on online hate with content found on Sri Lankan websites and social media. Without her, this report would simply not be.

Sanjana Hattotuwa
sanjanah@cpalanka.org

⁹ <http://sanjanah.wordpress.com/2013/02/01/anti-muslim-hate-online-in-post-war-sri-lanka/>

Executive Summary

A considerable amount of social media hate speech in Sri Lanka occurs on Facebook. The ability to like, share and comment on posts allows forums for supporters to engage, to plan rallies and other events and keep all similar posts in one place. It also allows admins of pages to remove and ban dissenting voices, allowing a greater degree of control than platforms such as Twitter. This phenomenon is not only relevant to Sri Lanka. According to the Umati Project in Kenya “only 3% of total hate speech comments collected by Umati originated on Twitter, while 90% were found on Facebook” (iHub Research & Ushahidi 2013: 5). Facebook also allows easier data collection on past events in comparison to Twitter. Twitter requires manual archiving in real time or sophisticated technical access to and analysis of its data streams, beyond the capacity and scope of this study (even though Twitter often acts in concert with content on Facebook to both flame violence and also dispel rumours)¹⁰. The degree to which hate speech exists on social media is often undetected by platform, domain or app owners (e.g. Facebook’s own hate speech monitoring mechanisms¹¹) due to the expression being predominantly in Sinhala (even the language used to annotate photos, illustrate videos or draw memes is predominantly if not exclusively Sinhala). This is why content in English that runs completely counter to Facebook’s policies around hate speech¹² finds free expression in Sinhala, only subject to scrutiny and compliance when reported by conscientious users. In highlighting some of these pages this analysis hopes to bring attention to the need for monitoring, further research and action against online hate speech in Sri Lanka, recognising at the same time that this is not a challenge limited to the country.

Ultimately, there is no technical solution to what is a socio-political problem. Sri Lanka’s culture of impunity and the breakdown in the rule of law is what affords the space for fascist groups like the *Bodu Bala Sena*, *Sinhala Ravaya* and *Ravana Balakaya* to say what they do and get away with it. In July this year, the Centre for Policy Alternatives (CPA) compiled a short brief to create awareness on the Constitutional and legal framework and available legal remedies regarding religious freedom and related issues, in light of the violence in Aluthgama¹³. What is evident is that even without new legislation looking at online domains and content, there are a range of legal remedies and frameworks to hold perpetrators of hate speech accountable for their violence, whether verbal or physical. The issue is not the non-existence of relevant legal frameworks, but their non-application or selective application.

If the trend around hate speech online is to be truly stemmed, media literacy programmes aimed at students, teachers, parents, lecturers, government officials, civil society, media consumers and citizens, in Sinhala and Tamil, over mobile friendly, freely accessible and engaging ways is required over the long-term. Episodic violence such as Aluthgama and Beruwela this year will test the ability of civil society to produce and promote counter-messaging to strategically and in a timely manner address the core submissions of

¹⁰ <http://www.firstpost.com/world/social-media-breaks-sl-medias-shameful-silence-on-aluthgama-riots-1572793.html>

¹¹ <https://www.facebook.com/notes/facebook-safety/controversial-harmful-and-hateful-speech-on-facebook/574430655911054>

¹² <https://www.facebook.com/communitystandards>

¹³ <http://www.cpalanka.org/constitutional-and-legal-framework-governing-religious-freedom-and-relatd-issues/>

hate speech online. This will also require civic media initiatives and citizen journalists to tie up with principled professional journalists working in mainstream media to fact check, debunk, engage and dispel rumours over social media in close to real time – vital skills that are aided by constantly improving digital media literacy (see for example resources from Amnesty International on social media content verification the Centre for Policy Alternatives has translated into Sinhala and Tamil, with a view to stemming rumours and misinformation online by those who are on social media platforms¹⁴).

Though there is no easy or prescribed solution, progressive thinking, proactive content production, strategic interventions and careful monitoring can identify and neutralise the wider harm online hate speech can, if unchecked and allowed to grow, sustain and strengthen.

Civility, tolerance and respect for diversity are as hard to find online as they are in Sri Lanka's mainstream party political framework even post-war. It would be a tragedy if the country's only remaining spaces to ideate, critical reflect and robustly debate – which are online - are taken over by hate-mongers, to the extent they are allowed to do so in the real world. So many in Sri Lanka, in various ways, resist violence, whether verbal or physical. The challenge is to strengthen their voices and efforts in light of what is a growing trend of hate speech production online, which though by no means easy, is also not an insurmountable one.

¹⁴ <http://groundviews.org/2014/08/08/verifying-user-generated-content-online-rough-guides-in-tamil-sinhala-and-english/>



The Centre for Policy Alternatives (CPA) is an independent, non-partisan organization that focuses primarily on issues of governance and conflict resolution. Formed in 1996 in the firm belief that the vital contribution of civil society to the public policy debate is in need of strengthening, CPA is committed to programmes of research and advocacy through which public policy is critiqued, alternatives identified and disseminated.

Address : 24/2 28th Lane, off Flower Road, Colombo 7, Sri Lanka
Telephone : +94 (11) 2565304/5/6
Fax : +94 (11) 4714460

Web : www.cpalanka.org, facebook.com/cpasl, twitter.com/cpasl
Email : info@cpalanka.org