biiss Papers 28

INFORMATION DISORDER IN THE INFORMATION AGE: ACTORS, TACTICS, AND IMPACTS IN SOUTH AND SOUTHEAST ASIA

Ayesha Binte Towhid

Bangladesh Institute of International and Strategic Studies (BIISS)

biiss papers

NUMBER 28 OCTOBER 2020

Bangladesh Institute of International and Strategic Studies (BIISS)
Dhaka

biiss papers

Number 28, October 2020

Published by
Bangladesh Institute of International and Strategic Studies (BIISS)
1/46, Old Elephant Road, Ramna
Dhaka-1000, Bangladesh.

Subscription Rate
BDT 150.00/USD 15.00 (Air Mail Charge Extra)

For correspondence please contact

Publication Officer
Bangladesh Institute of International and Strategic Studies (BIISS)
1/46, Old Elephant Road (West of Ramna Police Station)
Dhaka 1000, Bangladesh.
Phone: (880-2) PABX: 9353808, 9336287, 8315808, Ext.136
Fax: 8312625, e-mail: po@biiss.org, website: www.biiss.org

Printed by

GraphNet Limited
95, Naya Paltan, 1st Floor, Dhaka-1000, Bangladesh
Phone : 9354142, 9354133, e-mail: graphnet@gmail.com
Cell: 01715011303, website: www.graphnet.com

biiss papers
Number 28 October 2020

INFORMATION DISORDER IN THE INFORMATION AGE: ACTORS, TACTICS, AND IMPACTS IN SOUTH AND SOUTHEAST ASIA

Ayesha Binte Towhid

Ayesha Binte Towhid is Research Officer at Bangladesh Institute of International and Strategic Studies (BIISS). Her e-mail address is: ayesha@biiss.org

NOTE FOR THE CONTRIBUTORS

Original contributions (along with an abstract of 200-300 words) not published elsewhere may be submitted to the Chief Editor (Director General, BIISS)/Editor in duplicate, typed double-spaced, normally within about 30000 words. Footnotes should be placed at the bottom of the page following the styles given below:

For Books

Author, *Title*, Place of publication: Publisher, Year of publication, Page.

Example: one author

G. H. Johnson, *Harper's Connection*, Boston, USA: Penguin Books, 1998, p. 23.

Example: edited books

J. P. Forgas (ed.), *Feeling and Thinking: The Role of Affect in Social Cognition*, New York: Cambridge University Press, 2000, p.12.

For Chapters in Books

Author(s), "Title", in Author(s) (eds.), *Book Title*, Place of publication: Publisher, Year of publication, Page.

Example:

R. Macklin, "Conflicts of Interest", in T. L. Beauchamp and N. E. Bowie (eds.), *Ethical Theory and Business,* Englewood Cliffs, New Jersey: Prentice-Hall, 1983, pp. 240-246.

For Journal Articles

Author, "Article Title", *Journal Title*, Volume, No/Issue, Year of publication, Page number(s).

Example:

Sufia Khanom, "Gender Issues in Climate Change: Bangladesh Perspective", *BIISS Journal*, Vol. 30, No. 4, 2009, p. 450.

For Documents

Department/Agency, *Title*, Place of publication: Publisher, Year.

Example:

Department of HM Inspectorate of Constabulary, *Police Integrity, England, Wales and Northern Ireland: Securing and Maintaining Public Confidence*, London: Home Office Communication Directorate/HMIC, 1998.

NOTE FOR THE CONTRIBUTORS

For Newspaper
Author, "Article Title", *Newspaper Title*, Day Month Year.

Example:
G. M. Kabir, "Energy Crisis", *The Daily Star*, 15 December 2008.

For Paper Presented in Seminar/ Workshop/ Conference
Author, "Title of the Paper", paper presented in the Seminar/Workshop/Conference on *Title*, organized by ..., Place, on Day Month Year.

Example:
Roy Isbister, "Introduction to Illicit SALW Trafficking", paper presented in the Regional Conference on *Peace and Security in South Asia: Issues and Priorities for Regional Cooperation on Small Arms and Light Weapons Control,* organized by BIISS and Saferworld, Dhaka, on 08-09 November 2009.

For Web Document/Site
Author, "Title of Document", available at web address, accessed on Day Month Year.

Example:
G. H. Johnson, "Harper's Connection", available at http://www.m-q.edu.au/12, accessed on 12 March 2010.
If no author and title, give web address and access date only.

Example:
Available at http://www.mq.edu.au/12, accessed on 12 March 2010.

Tables, maps and diagrams should be placed in separate sheets. Contributors are also requested to enclose a brief biographical note and contact address in separate sheets. Scripts submitted for publication are not returned. For book review, two copies of book should be sent to the Chief Editor.

EDITORIAL BOARD

CONTACTS

Designation	Telephone (Office)	E-mail
Chairman	88-02-9347914	chairman@biiss.org
Director General	88-02-8312609	dgbiiss@biiss.org
Research Director-1	88-02-9331977	rd1@biiss.org
Research Director-2	88-02-8360198	mahfuz@biiss.org

biiss papers
Number 28 October 2020

CONTENTS

ABSTRACT

Access to abundant information affordably and instantaneously has been the most empowering feature of the 'Information Age'. Nonetheless, it results in rampant spread of disinformation, misinformation, malinformation, hate speech, defamatory remarks, and rumours by actors like troll farms, cyber troops, fake news syndicates, bot networks, hard-line religious groups, political parties, and even by some governments. This has given rise to an 'Information Disorder', particularly in cyberspace, which has been affecting different parts of the world in varied magnitude. South and Southeast Asia need particular focus in this regard as the countries have faced severe real-life impacts like vigilantism, mob killings, hate speech instigating attacks on minority communities, revenge attacks surrounding religious defamation, etc. Moreover, it has ramifications for democracy, law and order situation, and national security. In this backdrop, the paper undertakes an in-depth analysis of the nature and extent of information disorder in these two regions by studying the actors, tactics, and impacts. Both primary and secondary sources are used in the study. Using a contextualized framework, the paper identifies the major actors, their motivations, and the kind of contents created and disseminated by them. It finds that the underlying social, political, and economic factors, particularly the religious, ethnic, and racial fault lines play a key role in the process. Changes in the political environment and media landscape, technological transformation, and the business model of online platforms are also major factors. Although the problem is country-specific, there are transnational and regional implications as well. In this regard, the paper suggests some proactive, reactive, immediate, and long-term approaches that can be adopted by the countries individually, through regional collaboration, and by participating in global advocacy.

Keywords: Information Age, Information Disorder, Disinformation, Social Media, South Asia, Southeast Asia

ACRONYMS

ABT	Ansarullah Bangla Team
AI	Artificial Intelligence
AMRI	ASEAN Ministers Responsible for Information
ASEAN	Association of Southeast Asian Nations
BBC	British Broadcasting Corporation
BBS	Bodu Bala Sena
BEI	Bangladesh Enterprise Institute
BIMSTEC	Bay of Bengal Initiative for Multi-Sectoral Technical and Economic Cooperation
BJP	Bharatiya Janata Party
BTRC	Bangladesh Telecommunication Regulatory Commission
COVID-19	Coronavirus Disease
CENS	Centre of Excellence for National Security
CIA	Central Intelligence Agency
CSIS	Center of Strategic and International Studies
EU	European Union
FDA	Food and Drug Administration
GEC	Global Engagement Center
GIFCT	Global Internet Forum to Counter Terrorism
HuJI-B	Harkat-ul-Jihad-al-Islami-Bangladesh
ICT	Information and Communications Technology
IFCN	International Fact-Checking Network
INC	Indian National Congress
ISIS	Islamic States/Islamic States in Iraq and Syria
ISPR	Inter-Services Public Relations
ITU	International Telecommunication Union
JMB	Jama'atul Mujahideen Bangladesh
PCOO	Presidential Communications Operations Office
MCA	Muslim Cyber Army
OTI	Open Technology Institute
OII	Oxford Internet Institute
RSIS	S. Rajaratnam School of International Studies
SSRC	Social Science Research Council
TRS	The Real Singapore
UNB	United News of Bangladesh
UNESCO	United Nations Educational, Scientific and Cultural Organization
US	United States
UN	United Nations
UNICEF	United Nations Children's Fund
VTV	Vietnam Television
WHO	World Health Organization

Chapter 1

Introduction

The 'Information Age'[1] dwells in the dichotomy between information abundance on one hand, and information disorder on the other. Emerging through the digital revolution, political transitions, and economic developments in the mid-20th century,[2] and the meteoric rise of the internet, social media, and internet-enabled personal devices in the 21st century, the Information Age became the most significant period in the history of communication. It made available the highest volume of information ever accessible to people. It broke the traditional hierarchy and democratized the production and consumption of news and information. In this period, the tools of production shifted to those who were previously only the audience — now the audience could also become co-producers of content. Anyone from any part of the world could disseminate information and build audiences worldwide in seconds. However, the information environment which promised unprecedented opportunities to the people and indicated a golden era for citizen journalism, became subject to abuse by several actors like cyber troops[3], troll farms[4], bot networks[5], clickbaits[6], fake news syndicates[7], hard-line and extremist religious groups, communication agencies[8], and even some governments. Thus, the Information Age gradually turned into an age of information disorder where misinformation, disinformation, malinformation, rumour, and hate speech reigned.

While disinformation, propaganda, rumour and hate speech are not new, the powerful new tools of the Information and Communications Technology (ICT) driven Information Age made manipulation, fabrication, and amplification of contents easy, effective and hard-to-trace. The affordances of social networking technologies like algorithms, automation, and big data changed the scale, scope, and precision of how

[1] The term 'Information Age' is used to denote the present time which is witnessing a revolution in the use, flow and control of information. Joseph S. Nye Jr. and Robert O. Keohane extensively used the term to highlight the information-based world where the speed, technology and dramatic decreases in the cost of creating, processing and transmitting, and searching for information empowered a wide range of actors. This paper highlights these aspects of the Information Communications Technology dominated world and its impacts on society and state by focusing on the developments in the past two decades.

[2] Julian Birkinshaw, "Beyond the Information Age," Wired, available at https://www.wired.com/insights/2014/06/beyond-information-age/, accessed on 01 April 2020.

[3] Institutionalized groups of people tasked with the creation and dissemination of digital media content to shape and manipulate opinion on behalf of their affiliated political party or government agencies.

[4] Profiteering organizations consisting of people operating several online accounts usually to deceive the audience and disrupt conversations.

[5] Bits of code or software designed to interact with and mimic human users.

[6] Online entities which purposefully create misleading and attention-grabbing headlines to lead the audience to their sites for advertisement revenues. These can be harmful when the disputed content is related to matters of public interest.

[7] Organizations purposefully creating disinformation targeting social media platforms and online portals.

[8] Public relations agencies and consultancy firms that facilitate customized digital media campaigns.

information is transmitted in the present time.[9] Thus, disseminating manipulative and fabricated information reached a scale never witnessed before. All these gave rise to a polluted information ecosystem which is collectively labelled as the 'information disorder', where anyone starting from an individual to a state can abuse the technology to manufacture content for reasons starting from defamation of a person living next door to waging information operations in a country thousands of miles apart. While different countries have been experiencing the impact of this problem in varying magnitude over the past two decades, the need for holistically studying the topic started to gain prominence among scholars very recently, primarily after the 2016 United States (US) elections and Brexit referendum. In the following years, researches were carried out on related topics, but those mainly focused on information disorder in the democracies of the West. The digital landscapes in the Global South remained relatively underexplored, despite these regions being home to some of the fastest-growing digital markets in the world and having a plurality of political systems.[10] The academic works about these regions' information disorder are limited and mostly reactive to specific incidents, for example, Facebook's role during atrocities against Rohingyas in Myanmar and WhatsApp rumours inducing mob lynching in India. While these instances showed the extent of violence and harm that online speeches can lead to in these countries, the overall information disorder scenario in most countries of these regions remained comparatively understudied.

A post on social media may lead to online arguments, angry identity politics and polarization in other parts of the world, while emotionally charged rumours, panic over a perceived threat and posts hurting religious sentiments have often resulted in vigilantism and attacks on target communities in many countries of South and Southeast Asia.[11] On the one hand, these conflicts have been escalated by emotionally charged verbal cultures often rooted in long-standing ethnic, religious, and caste divisions present in many countries of the regions,[12] and on the other hand, this volatile situation also motivated several actors to bank on this opportunity to push forward their agenda through both manual and automated tactics. All these were facilitated by the rapid proliferation of ICT to a seemingly less prepared audience. While these have become a challenge for maintaining law and order, social cohesion, peace, and religious harmony in a country, in many cases it also has national security concerns and transnational implications among the countries in the two regions.

[9] Samantha Bradshaw and Philip N. Howard, *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation*, Oxford, UK: Oxford Internet Institute, 2019.

[10] Sahana Udupa, Iginio Gagliardone, Alexandra Deem and Laura Csuka, *Hate Speech, Information Disorder, and Conflict*, New York: Social Science Research Council, 2020.

[11] Amanda Taub and Max Fisher, "Where Countries Are Tinderboxes and Facebook Is a Match", *The New York Times*, 21 April 2018.

[12] Sahana Udupa et al., op. cit. p. 10.

The countries of South Asia and Southeast Asia are closely interlinked not only through geography but also through shared religions, ethnicities, languages, and cultures. Thus, it is seen that information campaigns in a particular country often have a ripple effect on the surrounding countries with similar communities. Moreover, many countries in these regions are heterogeneous and multicultural owing to the long history of migration between the two regions. During the colonial past and due to globalization in the last two decades, the two regions witnessed the migration of a large number of people.[13] These multicultural countries have witnessed both peaceful coexistence of the numerous groups and also conflicts among them in varying intensity grounded on these fault lines.[14] Information disorder targeting specific communities has added a new dimension to the identity conflicts in these countries. Moreover, dissemination of content across countries has never been easier, thus resulting in a real-time response from communities with similar identities in other countries. These show that the two regions are intricately connected as much in the virtual space as in the real world. For this reason, the paper attempts to study the information disorder scenario of South and Southeast Asia collectively to understand the several national, transnational, and regional dimensions of the problem.

In addition to the regional dynamics, in this hyper-connected world of the internet, the large digital media user base of these regions is also influenced by information campaigns unfolding at the international level, especially on topics that directly affect them, like the COVID-19 pandemic. The battle of narratives among China, Russia, and the US surrounding the pandemic also has an effect on how the crisis is perceived and addressed by the people here. This provides an international dimension to the problem. Thus, it can be seen that information disorder in these regions has emerged as a complex challenge with several dimensions that necessitates a holistic study. In this backdrop, this paper is an attempt to study information disorder in South and Southeast Asia by emphasizing on Bangladesh, India, Indonesia, Malaysia, Myanmar, Pakistan, Philippines, Singapore, Sri Lanka, Thailand and Vietnam. These countries have experienced significant impacts of information disorder on the societal and state level from organized actors which have drawn national and international attention. The impacts are well documented in these countries and the issue is acknowledged by different relevant stakeholders. In addition to these countries, disinformation and misinformation concerning Bhutan, Cambodia and Nepal are also briefly discussed. Insights of these fourteen countries allow to study the information disorder scenario of the two regions in a comprehensive manner.

[13] Kwen Fee Lian, Md Mizanur Rahman, Yabit bin Alas (eds.), *International Migration in Southeast Asia: Continuities and Discontinuities*, Singapore: Springer Singapore, 2016, p. 9.

[14] Aurel Croissant and Christoph Trinn, *Culture, Identity and Conflict in Asia and Southeast Asia*, Gütersloh: Bertelsmann Stiftung, 2009.

1.1 Research Objectives

There are three objectives of this paper. First, it aims to provide an analysis of why and how the information disorder in South and Southeast Asia developed over the last two decades by studying the actors, tactics, and impacts. It seeks to construct a conceptual framework based on the social and political realities of the countries in order to effectively study the information campaigns in these two regions. Second, it aims to bring forward the national, transnational, regional, and global dimensions of information disorder. Third, it attempts to suggest recommendations for effectively addressing the problem.

1.2 Research Questions

The paper has the following five research questions:

1. Who are the major actors responsible for information disorder and what are the motivations behind it?

2. What are the tactics involved in creating convincing content and what technology is deployed to disseminate it to the target audience?

3. What kind of impact the information disorder has on the society and state level?

4. What are the transnational, regional and international dimensions of the problem?

5. What can be effective ways of addressing this growing challenge?

1.3 Research Methodology

This paper is qualitative in nature and is based on both primary and secondary resources. It involved in-depth interviews consisting of academics, researchers, data scientists, media practitioners, fact-checkers and security analysts. The interviewees were selected based on their academic and professional association with the topic. The interviews were taken using a semi-structured questionnaire. The secondary data is collected from academic literature, reports of fact-checking organizations, policy briefs and publications of internet institutes, reports and disclosures of social media platforms, the study of digital labs, documents of international and regional organizations, government documents and the author's own analysis of digital media content on selected incidents in the past three years. The study also includes content analysis surrounding the COVID-19 pandemic but as it is an evolving issue, the analysis here is limited within the period from January to May 2020.

1.4 Limitations of the Study

This study has not been able to put equal emphasis on all countries of the two regions due to the lack of literature and constructive reporting of events in a few countries. While the study on both regions provides important insights on the transnational and regional dimension of the problem, it also risks generalization in few instances among the countries of this large and diverse sample size.

Following the introduction, the paper has been organized as follows. Chapter 2 develops a conceptual framework for understanding information disorder in the context of South Asia and Southeast Asia by defining different terms associated with it. Chapter 3 presents an overview of the major actors in the two regions. Chapter 4 analyzes the tactics used by the actors and the technological aspects of the online platforms. Chapter 5 studies the impacts of information disorder with relevant examples. Chapter 6 draws a pattern of the overall information disorder scenario in the two regions and brings forward the transnational, regional and international dimensions of the problem. Chapter 7 suggests way forward and Chapter 8 concludes the paper.

Chapter 2

Conceptual Framework

Information disorder is a complex phenomenon which is closely interlinked with the changing political and media landscape, monetization of content, incentives to manipulate, and the technological attributes of the Information Age. Before associating this phenomenon in the context of South and Southeast Asia, it is important to understand how this topic came into the global discussion. For this, it is important to trace back the political and economic developments in the mid to late 20th century during which the media landscape began to undergo significant changes and expansion. Media started to become deregulated in many countries, multi-channel television grew, the number of newspapers rose, first-generation internet access via dial-up modems started to spread and the number of options available to audiences and advertisers continued to expand between the 1980s and 1990s.[15] This change of media landscape accelerated in the first decade of the 21st century as the pattern of content creation, distribution and consumption witnessed even more radical shifts. The rise of digital technologies brought profound changes in how people communicate, interact, and learn about the world. While on one side there was an emergence of dominant search engines, the explosive growth of social media sites, and the spread of mobile web access,[16] on the other side, there was a decline in printed newspaper circulation and decreasing popularity of television as the key source of information and entertainment.

In this changing environment, online media platforms began to take the center stage. People increasingly became dependent on online media platforms to meet their information, communication and entertainment needs. Several factors contributed to this process like the penetration of the internet to all parts of the country including remote areas, availability of low-cost smart devices, affordable mobile data packages, on-demand access and the urge to receive and send content in real-time as it happens. This also led to the democratization of the media system as it removed barriers to publication. Anyone could produce content and share it with people anywhere in the world. However, the online platforms did not take into account how informed the users are, either as disseminators or consumers of information.[17] Also, until very recently, it did not differentiate who are the actors and what are their motivations for generating information. Thus, any kind of content could reach a global audience without undergoing the filters of traditional media houses. Also, nothing on the internet is lost, so these kinds of contents continued to float in the information ecosystem and reappeared in front of the audience constantly and in intervals. The volume of such contents became so vast that it

[15] Rasmus Kleis Nielsen, *Ten Years that Shook the Media World*, Oxford, UK: The Reuters Institute for the Study of Journalism, 2012.
[16] Ibid.
[17] Rituparna Banerjee, *Information Crises: The EU's Response and the South Asian Digital Media Landscape*, Brussels: South Asia Democratic Forum, 2020.

often diluted credible voices and changed people's perception of news and information. This ultimately resulted in a chaotic online environment which became popularly termed as 'information disorder'.

When it comes to information disorder, there are different concepts regarding the elements which pollute the information ecosystem. One of the most relatable terms for general people at present is fake news. Google Trends map shows that people mainly began searching for the term in the second half of 2016 and over the next years it witnessed exponential growth.[18] Although the term is very popular, it does not cover the entire gamut of information disorder. It leaves out genuine content that is shared out of context, manipulated content, and misleading statistics and visuals. Also, the term has been highly politicized to describe reporting that one does not agree with. However, the term is still widely used in the reporting and government statements in a few countries, so it will be used in specific cases in this paper.

Some of the other popular terms associated with information disorder include junk news, misinformation, propaganda, rumours, conspiracy theories, hate speech, defamatory speech, etc. Different stakeholders like think tanks, media practitioners, digital intelligence community, social media platforms, and governments have used different definitions in describing these elements. However, there are several challenges in defining the terms in a way which is acceptable by all the countries and reflects the entirety of the problem. Even within the countries, there lies several debates surrounding the terms as these can often be interpreted differently by different groups of people. This indicates that the problem is far more complex and has several grey areas beyond labelling a piece of news plainly as true or false. Thus, the definitional challenge has been one of the core issues of studying information disorder. However, despite the differences in opinion regarding definitions, a consensus is building up nationally and globally regarding the need to take measures to tackle the severe consequences that many countries of the world are facing right now due to the widespread of information disorder. In this regard, the paper explores some of the existing frameworks and then presents an amended version to effectively understand the problems prevalent in the countries of South and Southeast Asia.

First Draft's study on 'Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making' and 'First Draft's Essential Guide to Understanding Information Disorder' is widely followed in academic literature and fact-checking training.[19] First Draft's framework defines information disorder as a combined phenomenon of three main types of information, i.e., disinformation, misinformation and malinformation.

[18] Google Trend Map of the term Fake News, available at https://trends.google.com/trends/explore?date=today%205-y&q=fake%20news, accessed on 02 April 2020.

[19] First Draft is an international partner network of newsrooms, universities, platforms, and civil society organizations who develops tools and techniques for studying the information ecosystem of the world.

2.1 Disinformation

According to First Draft, "Disinformation is content that is intentionally false and designed to cause harm. It is motivated by three distinct factors: to make money; to have political influence, either foreign or domestic; or to cause trouble for the sake of it."[20] The United Nations Educational, Scientific and Cultural Organization's (UNESCO) Handbook for Journalism Education and Training defines disinformation as "deliberate (often orchestrated) attempts to confuse or manipulate people through delivering dishonest information to them...disinformation is particularly dangerous because it is frequently organized, well resourced, and reinforced by automated technology."[21] Disinformation also includes deceptive advertising in politics and government propaganda.[22] Oxford Internet Institute combined the two concepts, i.e., propaganda and automated technology and used the term 'Computational Propaganda' which refers to the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks.[23] Another group of scholars including Yochai Benkler, Robert Faris, and Hal Roberts terms this phenomenon as 'Network Propaganda' and denotes disinformation as a subset of propaganda that includes, "dissemination of explicitly false or misleading information".[24] It also incorporates "propaganda whose source or content is purely false, as well as propaganda whose source and content is more subtly masked and manipulated to appear other than what it is."[25] Thus disinformation can be broadly defined as an intentional or deliberate attempt to manipulate or influence people using false, dishonest and misleading information and propaganda. Here the intention of the actor and the objectivity of the content is central. Disinformation is the most widely used typology of information disorder and disinformation campaigns are widely practiced in almost all parts of the world, including South and Southeast Asia.

2.2 Misinformation

As per the First Draft's framework, the second typology is misinformation. First Draft describes, "When disinformation is shared it often turns into misinformation. Misinformation also describes false content but the person sharing does not realize that it is false or misleading. Often a piece of disinformation is picked up by someone without realizing it is false, and shared with their networks, believing that they are helping."[26] The UNESCO's handbook refers to misinformation as, "misleading information created or

[20] Claire Wardle, *First Draft's Essential Guide to Understanding Information Disorder*, UK: First Draft, 2019.
[21] Cherilyn Ireton and Julie Posetti (eds.), *Journalism, 'Fake News' & Disinformation: Handbook for Journalism Education and Training*, Paris: UNESCO, 2018, p. 7.
[22] Don Fallis, "What Is Disinformation?", *Library Trends*, Volume 63, Number 3, Winter 2015, pp. 401-426.
[23] Samuel C. Woolley and Philip N. Howard, *Computational Propaganda Worldwide: Executive Summary*, Oxford, UK: Oxford Internet Institute, 2017.
[24] Yochai Benkler, Robert Faris, and Hal Roberts, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*, New York: Oxford University Press, 2018, p. 32.
[25] Ibid.
[26] Claire Wardle, op. cit, p. 8.

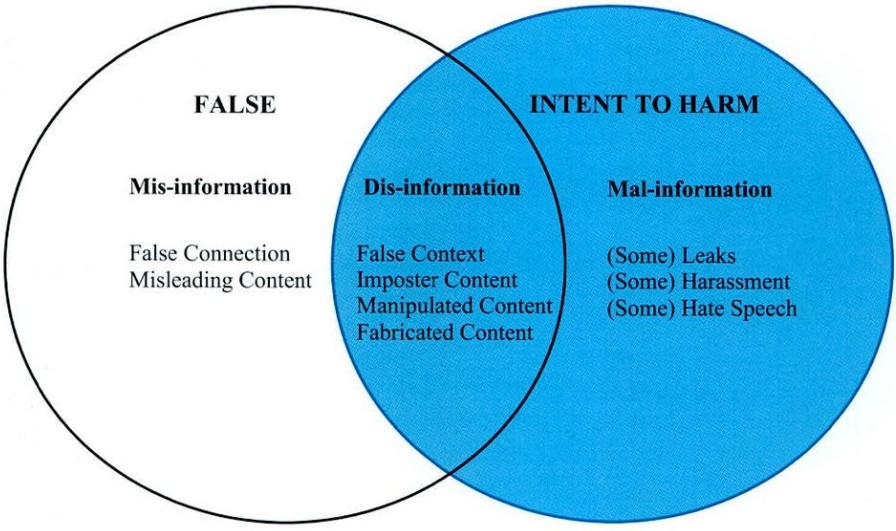
disseminated without manipulative or malicious intent."[27] The sharing of misinformation is driven by several motivational factors like the willingness to feel connected to their peers, be it the same political party, religion, race, or ethnic group or believers of certain issues. Many people also share misinformation because of their desire to be seen as a local 'expert' or 'first source' for local information.[28] Common people are the worst victims of this kind of information as they unknowingly amplify the reach of the disinformation and serve the purpose of the actors behind it, like 'useful idiots'.

2.3 Malinformation

The third category used by First Draft is malinformation which is described as, "genuine information that is shared with an intent to cause harm." Claire Wardle and Hossein Derakhshan define it as information "based on reality, but used to inflict harm on a person, organization or country".[29] These are messages with some truth but are created, produced, or distributed by actors who intend to harm rather than serve the public interest.[30] Malinformation is crucial in the present context. As social media companies have tightened up their ability to shut down fake accounts and changed their policies to be more aggressive against fake content, the actors of information disorder have learned that using genuine content reframed in new and misleading ways is less likely to get picked up by Artificial Intelligence (AI) systems and be deemed ineligible for fact-checking.[31] Therefore, many actors have adopted this technique of creating malinformation to ensure that their contents are widely circulated but not easily reported. While at times many actors deploy the combination of all three kinds of information campaigns across different platforms, it is important to understand the distinctions between them for analyzing the role of different actors involved in the process and challenging their tactics.

[27] Cherilyn Ireton and Julie Posetti, op. cit, p. 7.
[28] Shakuntala Banaji and Ram Bhat, *WhatsApp Vigilantes: An Exploration of Citizen Reception and Circulation of WhatsApp Misinformation Linked to Mob Violence in India*, London: Department of Media and Communications, The London School of Economics and Political Science, 2019.
[29] Claire Wardle and Hossein Derakhshan, "Thinking About 'Information Disorder': Formats of Misinformation, Disinformation, and Mal-information", in Cherilyn Ireton and Julie Posetti (eds.), *Journalism, 'Fake News' & Disinformation*, Paris: UNESCO, 2018, p. 46.
[30] Ibid.
[31] Claire Wardle, op. cit.

Figure 2.1: First Draft's Examination of How Mis-, Dis-and Mal-information Intersect Around the Concepts of Falseness and Intent to Harm.[32]



Here, it can be seen that all three categories of information disorder are focused on the intention of the actor and the objectivity of the content in use. While all these kinds of information disorder are widely created and disseminated in countries of South and Southeast Asia, the information disorder scenario in these regions is a bit more complex and these three terms are not adequate for addressing it holistically. For example, hate speech is a critical problem in most countries in this part. Although First Draft's framework very briefly addressed hate speech and included some types of hate speech and harassment under the malinformation category[33], this topic demands more focus in the context of South Asia and Southeast Asia. Also, while these three categories are helpful to understand the motive behind the actors, it does not adequately address the consequences.[34] In this regard, the paper further categorizes information disorder based on the action that it can lead to and its consequences or impacts.

[32] Framework created by Claire Wardle and Hossein Derakhshan, 2018, op. cit., p. 44.

[33] Claire Wardle and Hossein Derakhshan, *Information Disorder Toward an Interdisciplinary Framework for Research and Policymaking*, Strasbourg: Council of Europe, 2017.

[34] The author acknowledges the contribution of Din M. Sumon Rahman, Professor, Department of Media Studies & Journalism, University of Liberal Arts Bangladesh, in rethinking the existing framework.

2.4 Hate Speech

The first element that needs special attention in the context of South and Southeast Asia is hate speech. In many instances, it is seen that hate speech can threaten a community, risk violence against the community, trigger attacks, and also escalate ongoing conflicts. But hate speech is hard to define as there is no international legal definition of it. Different countries perceive it differently based on the socio-cultural and political context. Online platforms also have their own set of definitions in their policies and community standards to determine what kind of content will be allowed to stay on the platforms. But it is difficult to find a commonly accepted definition that can be consistently used. However, there are several definitions that give an essence of what hate speech comprises of. In this regard, the United Nations Strategy and Plan of Action on Hate Speech can be used to understand the broad concept. The document defines hate speech as "any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor. This is often rooted in, and generates intolerance and hatred and, in certain contexts, can be demeaning and divisive."[35] While hate speech can be disseminated through various media and have different consequences, this paper focuses on the online origin of the problem and actions that it can incite widely. In this regard, the functionality of hate speech can be seen through an in-group and out-group effect. In the Social Science Research Council (SSRC)'s Research Review on 'Hate Speech, Information Disorder and Conflict', hate speech is viewed as content that produces an out-group effect where, "target communities are seen as a threat to the safety and values of communities hate speakers claim to represent... it also has an in-group function in terms of recruiting and socializing new members and strengthening in-group memory. By exchanging and repeating hateful expressions targeting an out-group, group solidarities are built through rhetorical means and memory politics."[36]

To further stress the severity of the issue and highlight the real-life consequences of hate speech, academics have preferred to introduce the term 'Dangerous Speech'. The term was coined by Susan Benesch, Faculty Associate, Berkman Klein Center for Internet & Society at Harvard. 'Dangerous Speech' aims to provide a narrower but extreme version of hate speech based on its capacity of inspiring harms like mass violence. The guideline of the Dangerous Speech Project defines it as, "any form of expression (e.g., speech, text, or images) that can increase the risk that its audience (the in-group) will condone or commit violence against members of another group (the out-group)."[37]

---

[35] United Nations, *UN Strategy and Plan of Action on Hate Speech*, New York: United Nations Headquarters, 2019, p. 2.

[36] Perry, Barbara, *In the Name of Hate: Understanding Hate Crimes*, New York: Routledge, 2001 in Sahana Udupa et al., *Hate Speech, Information Disorder, and Conflict*, New York: Social Science Research Council, 2020.

[37] Susan Benesch, Cathy Buerger, Tonei Glavinic, and Sean Manion, "Dangerous Speech: A Practical Guide", available at https://dangerousspeech.org/guide/, accessed on 25 June 2020.

Here, the emphasis is put on the social, historical, and cultural contexts in which speech was made or disseminated to assess its possible impacts. While the term 'Dangerous Speech' is very useful for academics in assessing the potential threat of violence against a community, it is a relatively new term and is not relatable to many people. Also, the legal frameworks and reporting of incidents in most countries refer to such acts as hate speech. So, this paper uses the term hate speech but it emphasizes the severity of action that dangerous speech tried to indicate, for example, speeches to provoke exclusion, sow hate, fear, distrust, discord among the people, and incite real-life violence on target communities. In the case of South and Southeast Asia, the target communities of such kinds of hate speech are mostly ethnic and religious minority groups.

2.5 Religious Defamation

The second element of this impact-oriented category is religious defamation. Defamatory content can be hurtful for any individual who is targeted. However, in the context of South Asia, it is seen that a defamatory content shared online hurting religious sentiment of a particular community can result in a backlash from the target community in the form of revenge attacks which often turn into deadly clashes resulting in the destruction of property, religiously significant establishments and loss of lives. Thus, defamation also needs to be separately addressed. But defamation is also a very broad term and is subjected to several interpretations. In this case, the paper only focuses on defamatory content hurting religious sentiment which is disseminated through online platforms using a fabricated content or fake account to frame someone. Defamation from authentic sources is outside the purview of this study. Thus, the proposed framework of this paper only includes religious defamation using fabricated content or framed account as part of the information disorder.

2.6 Rumours Instigating Attacks and Vigilantism

Rumour is an age-old practice. But the online platforms have created the scope for easy, untraceable, and wide circulation of rumour by anyone with access to a smartphone. While many rumours can be harmless, in the region of South and Southeast Asia, it is seen that emotionally charged rumours that prey on the fear of people, for example, child kidnapping or forced sterilization[38], have often given rise to vigilante justice like lynching and mob attacks. In several instances, it was seen that people believed in the alleged rumours sent to them through peer-to-peer messaging services or social media posts and decided to take actions without relying on the law enforcement agency. This problem has been intensifying in many countries of both the regions in the past few years and thus requires special attention in studying how information disorder can inflict real-

[38] Rumour of forced sterilization by the Muslims to reduce the Sinhalese population became viral across social media platforms and messaging services in Sri Lanka leading to mobs setting fire to Muslim businesses. This will be elaborately discussed in Section 5.1 of the paper.

life damage. All these three elements give a new dimension to understanding information disorder in countries of South and Southeast Asia.

2.7 Association of Actors

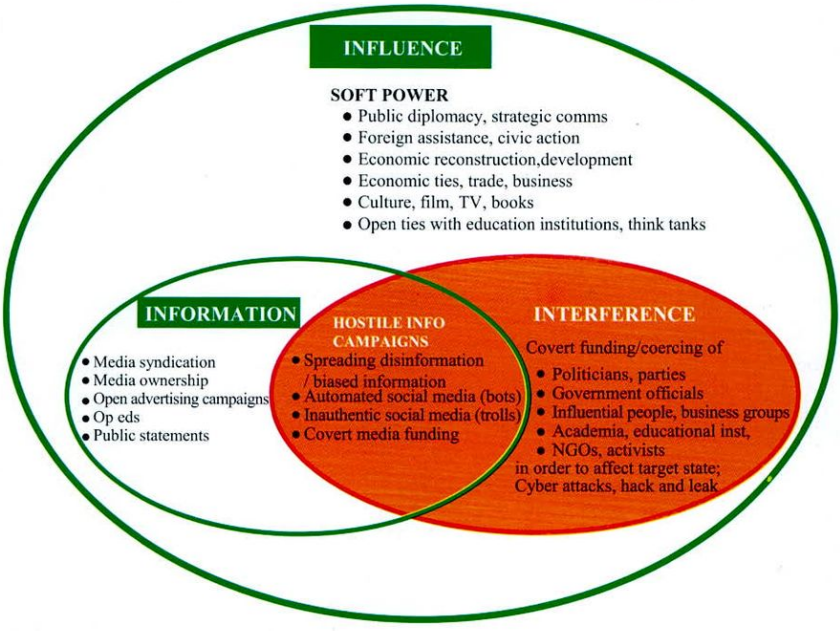
While the intention of the actors and the action it can trigger are understood, it is also necessary to trace the origin and association of the actors. In the present time, it is seen that information disorder can be attributed to both domestic and foreign sources. But the security implications and responses are fundamentally different between the two.

Domestic sources involve actors that are citizens of the country and operate within the country. They can be affiliated with the political groups, government, hard-line religious groups, troll farms, online media houses, etc. They can also be working independently or through informal groups in the country. In some cases, it is seen that citizens stationed outside the state are also engaged in disinformation. In this paper, they are also considered within the domestic category.

When it comes to foreign sources, the actors can be two kinds. First, they can be profiteering actors like troll farms of a foreign country engaged in information disorder campaigns against a target country. Second, the actors can be associated with a foreign state. Based on the scale of information campaigns from a foreign state, several terms are associated with it like, information warfare, information operation, psychological warfare, influence operation, malign influence operation, hybrid warfare, etc. However, on the international level, the definitions are even more nuanced. How countries perceive and respond to state-level campaigns often depend on power dynamics and bilateral relations. In this regard, the Centre of Excellence for National Security (CENS) of the S. Rajaratnam School of International Studies (RSIS) used the term 'Hostile Information Campaigns' which includes, "the spreading of disinformation or biased information in the defender state, spreading narratives by traditional media (such as newspapers) through proxies, or under covert identities and, carrying out these activities using automated social media accounts (bots) or inauthentic social media accounts (trolls) to create coordinated campaigns, often disguised as local opinions."[39] Acknowledging that the definitions can be fluid with several grey areas, and what is condemned as foreign interference by one nation may not be regarded as interference by another, RSIS proposed a framework for distinguishing between foreign interference, foreign influence, and hostile information campaigns.[40]

[39] Muhammad Faizal Bin Abdul Rahman, Gulizar Haciyakupoglu, Jennifer Yang Hui, Dymples Leong, Teo Yi-Ling and Benjamin Ang, *Countermeasures against Foreign Interference*, Singapore: S. Rajaratnam School of International Studies, Nanyang Technological University, 2020, p. 6.
[40] Ibid.

Figure 2.2: RSIS's Framework for Understanding the Relationship between Foreign Interference, Foreign Influence, and Hostile Information Campaigns.[41]



In light of the definition of Hostile Information Campaigns by RSIS and the illustration above, it is seen that a hostile information campaign is an intersection between interference and information and can be conducted through both traditional and online media. But as this paper focuses primarily on online platforms, only the online activities of such campaigns are examined for the regions of South and Southeast Asia.

2.8 Conceptual Framework of Information Disorder in South and Southeast Asia

Based on these discussions of the several forms of information disorder prevalent in the online platforms in South and Southeast Asia, the paper developed a conceptual framework that incorporates three major fields:

Intention of the actor and objectivity of content

This field involves disinformation, misinformation, and, malinformation. Here the focus is on the intention of the actor and the kind of content shared. Intention can be

[41] Muhammad Faizal Bin Abdul Rahman Gulizar Haciyakupoglu, Benjamin Ang, Dymples Leong, Jennifer Yang and Teo Yi-Ling, *Cases of Foreign Interference in Asia*, Singapore: S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University, 2020, p. 8.

of different kinds, for example, disinformation and malinformation have malicious intent and are usually shared by organized actors. On the other side, misinformation is usually shared by general citizens and may not have malicious intent but it largely contributes to the wide circulation of information disorder.

Objectivity of the content is also crucial. While it is acknowledged that there are grey areas surrounding it, it is seen that disinformation and misinformation shared online constitute false, manipulative and misleading information and, propaganda which can be mostly identified. On the other hand, malinformation may be based on genuine information but usually reframed in misleading ways and circulated with the intention to cause harm.

Action and scale of impact induced by information disorder

This field addresses online contents like hate speech, rumours and defamation through fabricated content or framed social media account. Such content needs special attention as it can result in large-scale impacts. These contents have the potential to trigger actions like violence against targeted communities or escalate an ongoing conflict. It can also lead to mob killings, vigilantism, and revenge attacks resulting in deadly clashes.

Association of the actors

The association of the actors can be of two kinds. First is the domestic category which involves citizens residing within the country or citizens living outside the state but engaged in information disorder campaigns targeting their respective country. Second, actors can be associated with foreign entities. In this framework, the foreign sources of information disorder would broadly include hostile information campaigns from a foreign state or state-sponsored agencies and, from foreign non-state actors like troll farms that can act on behalf of a foreign client and disguise themselves as local opinions.

While the term information disorder is collectively used to address the entire phenomenon, it can also be used to indicate any particular field of this framework based on the scale of its impact. This framework, illustrated in Figure 2.3, is used in the paper to comprehensively study the information disorder in South and Southeast Asia.

disorder to an entirely new scale.[43] Reports reveal that Macedonia was home to at least 100 pro-Donald Trump websites, many of them filled with sensationalist and utterly fake news and the people behind it made handsome profits in the process.[44] Such tactics are also widely practised in South and Southeast Asia, in fact on a much larger scale. Availability of low-cost manpower coupled with previous experience of working in clickbait farms, many countries in these regions have emerged as the hub of profit-making troll farms which offer their services to both domestic and international clients.

The Philippines is reported to be home to a large number of trolls and click farms. Over the years, the country has reportedly become home to hundreds of troll farms which manufacture any kind of fake content and false narratives depending on the client's wish. The industry is very vast and there appears to be services for all kinds of clients, like, trolls for companies, trolls for celebrities, trolls for liberal opposition politicians, and the government and even trolls for trolling trolls.[45] It is perceived that using the same young, educated, English-speaking workforce that made the Philippines a global call centre and content moderation hub, the troll farms in the Philippines are becoming a sought after destination for corporate and political campaigns worldwide.[46] While this brings short-term benefits to a few groups of people in the country, in the long-term it indicates the rise of an alarming level of information disorder in the country and the region.

Indonesia is another country where commercial disinformation actors are rapidly growing. Locally known as political 'buzzers' and fake-news syndicates, these agencies are responsible for producing fake news and are increasingly drawing politicians to take their services for defending their campaigns and/or spread misinformation or disinformation about their opponents on social media.[47] One of the most well-known buzzers Jonriah Ukur Ginting, popularly known as Jonru, had 1.47 million followers on Facebook, almost 100,000 on Twitter and more than 60,000 on Instagram and used his influence to promote disinformation and hate speech.[48] He was arrested in September 2017. Another notable producer of fake news and hate speech in Indonesia is Saracen, a syndicate of online content creators for hire. The administrators of this syndicate spread customized, divisive, and sectarian content using several thousand fake or hacked social

[43] Samanth Subramanian, "Inside the Macedonian Fake News Complex", *Wired*, available at https://www.wired.com/2017/02/veles-macedonia-fake-news/, accessed on 07 July 2020.

[44] Dan Tynan, "How Facebook powers money machines for obscure political 'news' sites", *The Guardian*, 24 August 2016.

[45] Shibani Mahtani and Regine Cabato, "Why crafty Internet trolls in the Philippines may be coming to a website near you", *Washington Post*, 26 July 2019.

[46] Ibid.

[47] Yenni Kwok, "Indonesia", in Masato Kajimoto and Samantha Stanley (eds.), *Information Disorder in Asia and the Pacific*, Hong Kong: Journalism and Media Studies Centre, The University of Hong Kong, 2019, p. 10.

[48] "Police arrest Jonru Ginting on charges of 'spreading hatred'", BBC Indonesia, available at https://www.bbc.com/indonesia/indonesia-41438278, accessed on 04 June 2020.

media accounts.[49] The Indonesian police uncovered the existence of Saracen in 2016 and arrested three of its members for spreading incendiary material online through social media.[50] However, this group kept functioning. In January 2019, Facebook removed 207 Facebook Pages, 800 Facebook accounts, 546 Facebook Groups, and 208 Instagram accounts linked to Saracen Group for engaging in 'coordinated inauthentic behaviour' on Facebook.[51] Although these financially-motivated actors might not always have a political agenda, their actions have a lasting impact on the political environment of a country.

While many studies have been carried out on the troll farms and fake news syndicates of the Philippines and Indonesia due to their global reach, there are limited studies and publicly available documents of similar actors in other countries of the regions. Although the troll farms and fake news syndicates may not operate in such magnitude or are not as organized, in the in-depth interviews for this paper, experts estimated that such actors are present and widely functional in other countries too. This can be understood by reflecting on the functions of similar agencies like Click Farms. Click Farms are largely functional in many countries of the region. For example, in Bangladesh, the operation of such companies can be traced back to 2012.[52] One effective tactic of information disorder is to use repurposed spam or marketing accounts for the dissemination disinformation and amplifying contents. The study of the datasets of recent global events showed that such accounts are readily and cheaply available for purchase from resellers.[53] So, it is predicted that such actors have evolved with the trend and engaged in information disorder activities as well.

3.2 Cyber Troops

Cyber troops can be loosely defined as groups of people with knowledge and skills of digital content creation, online engagement, and digital media advertisements. As digital media became an inevitable part of both civil and political lives, governments and political parties involved cyber troops in their media strategies to conduct public

[49] Yenni Kwok, op. cit.

[50] "Facebook takes down hundreds of Indonesian accounts linked to fake news syndicate", *Reuters*, 01 February 2019.

[51] Nathaniel Gleicher, "Taking Down Coordinated Inauthentic Behavior in Indonesia", Facebook Newsroom, available at https://about.fb.com/news/2019/01/taking-down-coordinated-inauthentic-behavior-in-indonesia/, accessed on 04 June 2020. This term is used by Facebook to indicate user behaviours which violate the company's community standard like people misrepresenting themselves, using fake accounts, artificially boosting the popularity of the content, or engaging in behaviours designed to enable other violations of its policies.

[52] Charles Arthur, "How low-paid workers at 'click farms' create appearance of online popularity", *The Guardian*, 02 August 2013 and Indrani Basu, "How 'Click Farms' In Bangladesh Are The Real Face Of Online 'Popularity'", Huffingtonpost, India Edition, available at https://www.huffingtonpost.in/2016/04/01/bangladesh-click-farms_n_9590000.html, accessed on 02 June 2020.

[53] Tom Uren, Elise Thomas and Jacob Wallis, *Tweeting through the Great Firewall*, Issues Paper Report No. 25, Canberra: The Australian Strategic Policy Institute, 2019.

diplomacy and manage their online presence. The size and permanency of the cyber troops varied depending on the media landscape and political environment of the countries. In some countries, the cyber troops are hired temporarily during elections or important political events, while in others, they are integrated into the media and communication landscape. But besides maintaining public relations of the governments and political parties, many cyber troops have often been deployed to control, censor, and shape conversation and information online.[54] In the 2019 Global Inventory of Organised Social Media Manipulation, Oxford Internet Institute (OII) presented evidence of organized social media manipulation campaigns in 70 countries in which at least one political party or government agency used social media to shape public attitudes domestically.[55] Out of these 70 countries examined, cyber troops of 52 countries reportedly created content such as memes, videos, fake news websites, or manipulated media in order to mislead users. OII's report from the previous year's shows that the number of countries undertaking such a form of computational propaganda was 48 in 2018 and 28 in 2017. That means, over the last two years, computation propaganda by governments and political parties using cyber troops increased by 150 per cent which shows the gravitas of the situation. The report also revealed the presence of cyber troops South and Southeast Asia which are briefly discussed in the following sub-sections.

3.2.1 *Cyber Troops Affiliated with Government Agencies*

Government affiliated cyber troops usually refer to people working for relevant agencies of the administration like communication or digital ministries or, in military-led campaigns. In OII's Global Inventory of Organised Social Media Manipulation, in the regions of South Asia and Southeast Asia, such actors were found in eight countries, i.e., Cambodia, Malaysia, Myanmar, Pakistan, Philippines, Sri Lanka, Thailand and Vietnam.[56] The scale of operation of cyber troops in these countries varies depending on the type of government and leadership.

The Philippines under the leadership of President Rodrigo Duterte has seen a drastic rise of misinformation, disinformation, and malinformation in both traditional and digital media. Such disinformation campaigns can be traced back to the 2016 presidential election during which a sophisticated information strategy was initiated by deploying a network of hyper-partisan influential bloggers, trolls, and bots.[57] It was reported that these people produced spin and spurious reports in different forms and on various platforms to influence voters during the election period.[58] After winning the election and taking

[54] Samantha Bradshaw and Philip N. Howard, 2019, op. cit.
[55] Ibid.
[56] Ibid.
[57] Maria A. Ressa, "Propaganda War: Weaponizing the Internet", Rappler, available at https://www.rappler.com/nation/148007-propaganda-war-weaponizing-internet, accessed on 04 June 2020.
[58] Yvonne T. Chua and Ma. Diosa Labiste, "The Philippines", in Masato Kajimoto and Samantha Stanley (eds.), op. cit. p. 25.

office, selected cyber troops from the election campaign were reportedly appointed in Presidential Communications Operations Office (PCOO) which indicated a continuation of the strategy at the state level.[59] This was particularly reflected surrounding the war on drugs. The cyber troops reportedly shared disinformation and malinformation by promoting distorted reports and statistics about the success of the war on drugs and used out of context images to justify actions.[60] The disinformation campaign also aimed at mass reporting and shutting down the social media accounts of opponents and critics of the government policies.[61]

In Vietnam, cyber troops have a formalized appointment. They are a part of a military cyber warfare unit called Force 47 and comprise of as many as 10,000 members.[62] Although the unit is responsible for countering 'wrong views' on the internet, they are also alleged for conducting online propaganda and censoring social media.[63] Such acts of cyber troops is a challenge for the information ecosystem.

Military operated computational propaganda was also reported in Myanmar. The Myanmar military has been responsible for persecuting minorities in the country through a series of oppressive acts. The military has been accused of using social media to propagate hatred against selected communities. On 28 August 2018, Facebook removed 425 Facebook pages, 17 Facebook groups, 135 Facebook accounts and, 15 Instagram accounts linked to the Myanmar military for engaging in 'coordinated inauthentic behaviour' on Facebook and Instagram.[64] On 15 October 2018, Facebook further took down 13 pages and 10 accounts linked to the Myanmar military under its misrepresentation policy.[65] Facebook's investigation found the Tatmadaw used seemingly independent news and opinion pages to covertly push the messages of the Myanmar military.[66] Here, deceptive measures were deployed to conduct disinformation campaigns by exploiting the features of digital platforms. There are also allegations that the military might have directly used social media platforms to incite violence against the Rohingyas. As part of the International Court of Justice case brought by Gambia accusing Myanmar of genocide against the Rohingya, a request was filed in June 2020 with the US District Court for the District of Columbia to force Facebook to release "all documents and communications produced, drafted, posted or published on the Facebook page".[67] It is expected that these

[59] Ibid, p. 26.
[60] Ibid.
[61] Miguel Syjuco, "Fake News Floods the Philippines", *The New York Times*, 24 October 2017.
[62] "Vietnam unveils 10,000-strong cyber unit to combat 'wrong views'", *Reuters*, 26 December 2017.
[63] Reporters without Borders, "Vietnam's 'Cyber-troop' Announcement Fuels Concern about Troll Armies", available at https://rsf.org/en/news/vietnams-cyber-troop-announcement-fuels-concern-about-troll-armies, accessed on 20 April 2020.
[64] Facebook Newsroom, "Removing Myanmar Military Officials From Facebook", available at https://about.fb.com/news/2018/08/removing-myanmar-officials/, accessed on 02 June 2020.
[65] Ibid.
[66] Hannah Ellis-Petersen, "Facebook removes accounts associated with Myanmar military", *The Guardian*, 27 August 2018.
[67] "U.S. court asked to force Facebook to release Myanmar officials' data for genocide case", *Reuters*, 10 June

communications materials of the Myanmar military officials and police forces including Min Aung Hlaing, commander-in-chief of Myanmar's armed forces might have evidence of genocidal intent against the community.[68] This is a significant revelation about the abuse of social media platforms by the actors of information disorder.

3.2.2 *Cyber Troops of Politicians and Political Parties*

Understanding the role of cyber troops appointed by politicians and political parties is complicated as there can be a thin line between information and propaganda in politics. Politicians and political parties across the world extensively use digital platforms as part of their campaign strategy. They use the power of the platforms to build grassroots support and spread their messages in crucial times like elections. However, besides promoting legitimate political messages, the cyber troops of many political parties have resorted to using disinformation. They are held responsible for intentionally spreading disinformation to forward their agenda, suppress political opponents and artificially inflate the number of followers, likes, shares or retweets using fake accounts, thus creating an illusion of grassroots support and popularity.[69] The existence of such cyber troops is very prominent in South and Southeast Asia. The 2019 Global Inventory of Organised Social Media Manipulation reported six countries namely, India, Indonesia, Malaysia, Pakistan, Philippines, and Sri Lanka for engaging in such activities.[70] The case of India and the Philippines are briefly discussed here to give an insight into the structure and methods used by the cyber troops.

In India, the leading political parties are reported to have deployed a large number of cyber troops for political campaigning on social media. Both the Bharatiya Janata Party (BJP) and Indian National Congress (INC) are known to have 'IT cells' since the early days of social media. However, many a time these groups have reportedly used automation, trolling and disinformation techniques.[71] The encrypted peer-to-peer messaging service of WhatsApp is the most used tool in the digital campaigns in the country, particularly during elections. For example, during the 2018 Karnataka state elections, it has been reported that at least 50,000 WhatsApp groups were created by both the BJP and INC to spread their messages.[72] While many messages in these groups were regular campaign advertisements, many were reported as disinformation, malinformation and, messages to inflame sectarian tensions among the different religious communities.[73]

2020.

[68] Ibid.

[69] Samantha Bradshaw and Philip N. Howard, *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*, Oxford, UK: Oxford Internet Institute, 2017.

[70] Samantha Bradshaw and Philip N. Howard, 2019, op. cit.

[71] Samantha Bradshaw and Philip N. Howard, "Case Study: India" in Samantha Bradshaw and Philip N. Howard, 2019, op. cit.

[72] Vindu Goel, "In India, Facebook's WhatsApp Plays Central Role in Elections", *The New York Times*, 14 May 2018.

[73] Ibid.

In the 2019 parliamentary elections, it was reported that BJP's Delhi IT Cell has built a five-tier system to reach its messages to the grassroots stretching from the states to the districts, and then on to booths, and finally *mandals*[74], employing several thousand workers for disseminating political messaging in WhatsApp groups.[75] But once these groups are created, it is difficult to control and monitor what messages are disseminated in it. Thus, it often becomes a source of disinformation. Alongside these two leading political parties, other groups have also resorted to similar campaigns to a varying extent. In a highly polarized political environment of India, the spread of disinformation by cyber troops of political parties has often led to the spread of misinformation by common people due to hyper-connectivity and lack of media literacy, thus giving rise to a highly complex information disorder.

A similar structure of cyber troops is also seen in the Philippines. The country witnessed a multi-layered network of political party affiliated cyber troops responsible for spreading disinformation. In a recent study, media scholars Jonathan Corpus Ong and Jason Cabanes documented an architecture of networked disinformation where professionalized and hierarchized groups of political operators design disinformation campaigns, mobilize click armies, and execute innovative digital techniques.[76] According to the report, at the top tier are 'chief disinformation architects' or strategists who are often senior professionals with an advertising and public relations background hired by politicians to put together and implement the communication plans. The chief disinformation architects, in turn, assembled 'digital influencers' who maintain multiple social media accounts with about 50,000 to two million followers across Twitter and Facebook. The chief strategists also hired community-level fake account operators to repost and retweet the messages of digital influencers to create illusions of engagement. The community-level workers are usually fresh college graduates, politicians' administrative staff, and online freelance workers.[77] Political parties and candidates in the Philippines are reported to have tapped into these networks during the elections to manipulate and shape the opinion of voters.

3.3 Extremists and Hard-line Religious Groups

Use of online platforms by local and international terrorist organizations have been a major security concern for many countries in South Asia and Southeast Asia. The use of these

[74] Sub-district level category of administrative division in some parts of India.

[75] Amrita Madhukalya, "56-hour lag in Congress's reach on WhatsApp gives BJP the edge", *Hindustan Times*, available at https://www.hindustantimes.com/lok-sabha-elections/who-s-creating-whatsapp-buzz-this-election-season/story-NAewnMqSCqZlxtGlaelDhO.html, accessed on 02 June 2020.

[76] Jonathan Corpus Ong and Jason Vincent A. Cabanes, "Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines", available at http://newtontechfordev.com/wp-content/uploads/2018/02/Architects-of-Networked-Disinformation-Executive-Summary-Final.pdf, accessed on 05 June 2020.

[77] Ibid.

platforms by the Islamic States (ISIS) to spread propaganda and radicalize young people added an entirely new dimension to the abuse of this technology. Besides these organizations, there is a new form of online activity by extremists and hard-line religious groups.

Religious groups and its leaders hold a very respectable and influential position in societies of many countries of South and Southeast Asia. However, many hard-line religious leaders and groups are reported to have abused their power to preach radical views, hate speech, and sow hatred and division in society along religious lines. Increasingly, such actors have adopted new technologies and initiated social media-based campaigns which have far-reaching consequences. In South and Southeast Asia, the active role of these actors is visible in countries like Myanmar, Thailand, Indonesia, India, Sri Lanka, and Bangladesh. In each country, there are several actors of this category belonging to Buddhist, Hindu, and Islamist hardliner groups. Here it needs to be noted that these groups are addressed through different labels, for example, radical groups, ultra-nationalists, religious-nationalist, ethno-nationalist, extremists, etc. However, this paper focuses only on the online content shared by these groups. Some of the major activities of the hardliners of dominant religious groups in their respective countries in recent times are briefly discussed in this part.

The first name that comes up in this category is Ashin Wirathu, the Buddhist monk of Myanmar and his hard-line Buddhist group Ma Ba Tha. They are held responsible for promoting disinformation and hate speech against the Rohingyas through social media accounts and pages. Ashin Wirathu has reportedly used social media for years to disseminate anti-Muslim hate speech, jingoistic sermons, and virulent rumours about the Rohingyas in the country.[78] He amassed hundreds of thousands of followers to help circulate his incendiary sermons and videos. In June 2018, Facebook blacklisted and banned Ma Ba Tha and few extremist monks in addition to Wirathu from its platform for posting inflammatory content and stoking hatred towards the Rohingyas.[79]

Such actors are also active in Myanmar's neighbouring country Thailand. For example, a radical Buddhist monk named Phra Maha Aphichat Punnajanto was held responsible for initiating social media campaigns to propagate anti-Muslim hate speech in the country. He came into discussion in 2015 after he urged Buddhists across the country to burn down a mosque as retribution for every monk killed in the insurgency in southern Thailand.[80] His messages also included claims that a number of Thai Muslims were conspiring to 'seize Thailand' through various means.[81] Such hate speeches are a threat to religious harmony and social cohesion in a country.

[78] Laignee Barron, "Nationalist Monk Known as the 'Burmese bin Laden' Has Been Stopped From Spreading Hate on Facebook", *Time*, 28 February 2018.

[79] AFP, "Facebook blacklists Myanmar hardline Buddhist group", *The Straits Times*, available at https://www.straitstimes.com/asia/se-asia/facebook-blacklists-myanmar-hardline-buddhist-group,accessed on 14 July 2020.

[80] Panarat Thepgumpanat and Panu Wongcha-um, "Thai government takes action against monk over anti-Muslim views", *Reuters*, 21 September 2017.

[81] "Muslim leader demands probe of anti-Islam talk", Bangkok Post, available at https://www.bangkokpost.com/thailand/politics/760780/muslim-leader-demands-probe-of-anti-islam-talk, accessed on 01 July 2020.

Hard-line Buddhist groups using social media to advance the ethno-nationalist agenda is also present in Sri Lanka. Amith Weerasinghe's Mahason Balakaya and Buddhist monk Galagoda Atte Gnanasara's Bodu Bala Sena (BBS) are the two most prominent groups at present.[82] These groups have been responsible for amplifying hate speech against the minority communities, particularly Muslims, and sowing divisions in the society. Social media pages run by leaders and activists of these groups are perceived to be critical tools of their hate campaigns.

Anti-Muslim hate speeches and disinformation are also widely disseminated by the Hindu zealots in India in recent times.[83] A study by Equality Labs revealed that Islamophobic content is often the most violent, threatening, and gruesome of hate speeches in Facebook India.[84] Such contents involved Islamophobic slurs, the glorification of violence by alluding to previous incidents of violence against Muslims, anti-Rohingya posts, conspiracy theories, etc.[85] Alongside targeting the Muslim community, hate speech and disinformation against the Christians and selected castes like the Dalits are also widely disseminated. Besides Facebook, several other platforms like WhatsApp and Twitter are widely used in this process.

Another actor in this category is the cyber armies banking on religious identity. Such actors use online platforms to exploit the religious and ethnic fault lines and trigger hate and intolerance. The Muslim Cyber Army (MCA) of Indonesia is one such group. According to The Guardian's report, "The network is accused of spreading fake news and hate speech to inflame religious and ethnic schisms; fan paranoia around gay men and lesbians, alleged communists and Chinese people; and spread defamatory content to undermine the president."[86] The MCA network is reported to have several groups tasked for different activities, namely the Family MCA, the United MCA, the Legend MCA, Special Force MCA, Muslim Sniper, and MCA News Legend.[87] In 2018, Indonesian Police arrested six suspected members of this network for spreading fake news in the country.[88] According to the report of the cybercrime directorate at the National Police's Criminal Investigation Unit, the group actively uploaded negative content and fake news on social media which

[82] Zaheena Rasheed, "In Sri Lanka, hate speech and impunity fuel anti-Muslim violence", *Al Jazeera,* 13 March 2018.

[83] See more at Maya Mirchandani, *Digital Hatred, Real Violence: Majoritarian Radicalisation and Social Media in India*, ORF Occasional Paper 167, New Delhi: Observer Research Foundation, 2018.

[84] T. Soundararajan, A. Kumar, P. Nair and J. Greely, "Facebook India. Towards the Tipping Point of Violence: Caste and Religious Hate Speech", Equality Labs, 2019, available at https://www.equalitylabs.org/facebookindiareport, accessed on 08 July 2020.

[85] Ibid.

[86] Kate Lamb, "Muslim Cyber Army: a 'fake news' operation designed to derail Indonesia's leader", *The Guardian*, 13 March 2018

[87] Vincent Bevins, "Indonesian police arrest 14 suspected members of radical Islamist cyber network", *The Washington Post*, 01 March 2018.

[88] Sheany, "Police Arrest Core Members of 'Muslim Cyber Army'", Jakarta Globe, available at https://jakartaglobe.id/news/police-arrest-core-members-of-muslim-cyber-army/, accessed on 08 July 2020.

included defamation of individuals or groups, and material that could provoke racial, ethnic, religious or intergroup conflict.[89]

Extremist groups in Bangladesh have also made online platforms an integral part of their campaign. Social media networks were extensively used by extremist groups like the Hizb ut-Tahrir Bangladesh and Ansarullah Bangla Team (ABT) to disseminate propaganda and encourage individuals to oppose the state.[90] According to Bangladesh Enterprise Institute (BEI)'s Perception Survey, alongside Hizb ut-Tahrir and Ansarullah Bangla Team, other groups like Jama'atul Mujahideen Bangladesh (JMB), Harkat-ul-Jihad-al-Islami-Bangladesh (HuJI-B) and Hizbut Tawhid are also major actors in using online mediums to propagate their radical views.[91] Besides these extremist organizations, at present, many smaller hardliner groups and religious leaders with radical views are also seen to use social media to propagate hate messages targeting other religious groups in the country.

3.4 Volunteers

Volunteers are another important actor in information disorder. They actively collaborate with government or political parties to spread political ideology or pro-government messages. In many cases, volunteer groups are made up solely of youth organizations. While these groups of volunteers usually promote a positive image of the government or political party, they often resort to sharing disinformation. Such groups are visible in almost all countries of South and Southeast Asia.

In India, networks of volunteers disseminate sophisticated disinformation strategies across social media, responding in real-time to political developments. There are several scopes of engagement, for example, assist in WhatsApp Group Management, create content or engage on social media platforms like Facebook and Twitter.[92] These volunteers often post false and hyper-partisan information in defending the political party they are connected to. Such allegiance also seems to have political motives like getting noticed by the party IT cell or the party political leaders themselves. And in case they get noticed or followed by the IT cell or the political leaders, it somehow increases their social credit or social capital and often opens doors for them in real life.[93]

---

[89] Ibid.

[90] Iftekharul Bashar, "ISIS, AQIS and the Revival of Islamist Militancy in Bangladesh", *Counter Terrorist Trends and Analyses*, Vol. 7, No. 6, 2015, p. 20.

[91] Bangladesh Enterprise Institute, *The Role of the Media in Countering Radicalisation in Bangladesh*, Dhaka: Bangladesh Enterprise Institute, 2014.

[92] See more at Karnika Kohli, "Congress vs BJP: The curious case of trolls and politics", *The Times of India*, 11 October 2013.

[93] Based on an interview with Rudroneel Ghosh, Indian analyst and journalist, The Times of India on 22 May 2020.

In Sri Lanka, the volunteers propagate disinformation on behalf of political parties or religious and ethnic groups. They act like trolls and engage in amplification of content favourable to them and drowning out criticisms of their favoured party or group.[94] Similar activities are also reportedly visible in Bangladesh, Myanmar and the Philippines. But in most times, these are loosely coordinated and thus difficult to trace.

Volunteers are important actors in the political field as they actively disseminate disinformation in support of political parties or politicians but it insulates them from any direct association. While most of the time such acts are intentional, at times general citizens perform such acts to show allegiance and loyalty to the party or leader without realizing they are contributing to the information disorder.

3.5 Social Media Influencers and Paid Citizens

As actors of information disorder, paid citizens have a slightly different role in comparison to the volunteers. At times it is seen that political parties actively recruit citizens who are usually in prominent positions in the society or on digital platforms in order to share specific narratives advantageous to them in exchange for money or any other form of remuneration. Since these citizens are not officially affiliated with the political party, their voices can be used to help disseminate messages from a seemingly independent voice. Such actors are perceived to be present in many countries, but because the arrangement takes place on a personal level, it is difficult to trace and document.

3.6 Individuals and Groups

Individuals and groups working independently are one of the most crucial actors of information disorder. As already discussed, the digital platforms equip its users with powerful technologies to produce and disseminate information. Individual users can run independent blogs, channels, pages, or websites through which he/she can share disinformation, malinformation, and hate speech targeting the state, government, and religious groups. Such actors are visible in all countries. For example, in Bangladesh, such individuals and their associates are some of the major actors of disinformation. They often act from outside the country or in a covert manner within the country and amplify their contents through various techniques.

At times individuals with firm beliefs regarding a particular issue can also form groups among themselves and further push out content through collaborated content creation and sharing. While this strategy was immensely powerful for activists, now it is exploited by sections of the online media users. Besides creating ill-intended contents, individuals and groups also contribute to the information disorder cycle by sharing misinformation without realizing the harm they are doing to the information ecosystem.

[94] Based on an interview with a Sri Lankan media practitioner on 08 May 2020.

3.7 Online Media Outlets and Online Versions of Traditional Media

The advent of new media has given rise to several small to large scale content creators catering to the need of online news portals, websites, and blogs. But these online outlets do not always maintain a standard editorial policy and often initiate disinformation campaigns. The digital advertisement models based on the level of user engagement is one of the main factors for these content producers to create news which grabs quick attention rather than objective reporting. Such practice is seen in many parts of the world including South and Southeast Asia. For example, online news websites are often the purveyors of fake news in Indonesia and Vietnam. India has its fair share of websites that spread misinformation, disinformation, and propaganda. In Bangladesh, online news portals are also significant factors behind information disorder.

Here, it is seen that in some cases, the online news portals target general audiences throughout the country, but also in some cases such platforms target limited audiences in selected regions within the country.[95] Many of these portals are known for spreading misinformation and rumours. Because of the scale and nature of their operation, these platforms are often difficult to monitor by fact-checkers. But these small actors play an influential role among the local audience and misinformation shared through these platforms can have severe consequences. Thus, it is also important to pay attention to such actors.

While the actors of online platforms are under serious scrutiny for their role in information disorder, the traditional media outlets are also responsible to some extent. Over the past years, most traditional media outlets have created their online presence. But it is often seen that the online version of mainstream media outlets does not maintain a similar editorial standard as that of the print or broadcast version. In order to stay relevant and break stories first, they often resort to unsubstantiated news, clickbait headlines, and disinformation. This is frequently seen in the case of Bangladesh and India. Disinformation and malinformation coming from these mediums are very concerning as they can cause severe harm due to their wide reach and acceptability. Additionally, many newspapers and media channels pick up unverified social media posts and reproduce it in the mainstream media without sufficient scrutiny, further contributing to the information disorder.

The above discussions presented an overview of the main categories of actors responsible for information disorder in South and Southeast Asia. The motivations and nature of the operation of the actors showed how the information disorder developed in the countries of these regions. Apart from these major actors, the operation of other actors like strategic communication firms, public relations agencies, and civil society organizations are also visible in few countries on a small scale. All these portray the

[95] Based on an interview with Din M. Sumon Rahman on 26 June 2020.

wide range of actors starting from individuals to governments, and from volunteers to industrial trolls, who are responsible for conducting different forms of misinformation, disinformation, malinformation, rumour and hate speech campaigns through online platforms and messaging services in order to create an information disorder which suits their specific purposes starting from winning elections to orchestrating genocides on minorities.

Chapter 4

Tactics Adopted in Information Disorder

The actors of the information disorder adopt a series of tactics to generate content suitable for their agenda and then spread it to their target audiences. Over the years these tactics have evolved depending on the skills of the actors, technological advancement, and the features of digital platforms. Several steps are involved in the process. In this paper, the commonly used tactics of the twofold process, i.e., creation and dissemination have been discussed with relevant examples of South and Southeast Asia.

4.1 Content Creation

The first step involves the generation of content. The contents can be classified into several types. In this part, the outline of the First Draft is taken as a primary basis of content classification and then additions are made based on the popular practices in the two regions.

4.1.1 *Fabricated Content*

Fabricated content indicates that it is 100 per cent false and is designed to deceive and do harm. The actors of information disorder use this tactic in multiple ways like creating false images, audio and videos. At times these are very easy to detect, but in many cases, the level of fabrication is so intricate that it often bypasses the filters of social media companies and fact-checkers. Increasingly, powerful technological features like Artificial Intelligence and Deepfakes[96] are being used to create deceiving content. While much discussions surrounding it have been going on in the West, a similar tactic was also seen in the recently held Delhi election in India. On 07 February 2020, a day ahead of polling in Delhi for the Legislative Assembly, a series of videos of BJP politician Manoj Tiwari appeared on multiple WhatsApp groups. The videos showed Manoj Tiwari speaking against his political opponent Arvind Kejriwal in English and Haryanvi.[97] But analysis of the Indian fact-checking organization Boom revealed that those videos were never made, rather manufactured by AI-generated Deepfake technology. It needs to be noted that in a political environment where information disorder is rampant, creating such fabricated content which is hard for any regular viewer to detect is alarming.

[96] The term Deepfake refers to synthetic media where the speech and actions of an individual can be used to create a video of another person using deep learning. Previously the term was mostly associated with pornography, but now it is also used in political campaigns.

[97] Archis Chowdhury, "BJP's Deepfake Videos Create Row; EC Unclear How To Respond", Boom, available at https://www.boomlive.in/politics/bjps-deepfake-videos-create-row-ec-unclear-how-to-respond-6952, accessed on 05 April 2020.

4.1.2 *Imposter Content*

Genuine sources of information are often impersonated to push out convincing content at the cost of deceiving the audience. This tactic often involves the logo or name of credible mainstream media outlets. This tactic is widely used by actors like cyber troops and fake news syndicates. Recently in many countries including Bangladesh, India, and Pakistan, a post went viral in social media regarding the COVID-19 that used UNICEF as the source. People believed in the content as it used the name and logo of a trusted organization. But it was misinformation and the UNICEF authority later debunked issuing such a statement.[98] In India, Sri Lanka, and Bangladesh[99], government agencies were also impersonated to create misleading content regarding the withdrawal of lockdowns and government decisions regarding the situation.

Another tactic of imposter content is creating false Facebook pages or YouTube channels replicating authentic news sites. This is a common practice by selected actors. For example, in Bangladesh, Facebook took down "nine Facebook pages designed to mimic legitimate news outlets", which included one appearing like British Broadcasting Corporation (BBC)'s Bangla-language service and another replicating bdnews24.com, a popular online news portal in the country.[100] In Vietnam, imposter accounts of Vietnam Television (VTV) and fake VTV channels are widely viewed on YouTube. Imposter sites are common in the Philippines as well. It is difficult for average readers to distinguish between authentic and fake sites at first glance. While such sites are usually taken down or deactivated after being exposed, it is seen that the contents often remerge through new identities.

4.1.3 *False Context*

This tactic involves the use of genuine content but with false contextual information. This is a widely used tactic in times of natural disasters or conflicts by most actors. For example, in Bangladesh, a photo of a man engulfed in flames was widely shared in social media attributing to the atrocities on Rohingyas in Myanmar. However, the photo was originally of a Tibetan activist who set himself on fire and ran through a demonstration in New Delhi ahead of the Chinese President's visit to India on 26 March 2012.[101] Many such contents were shared regarding the Rohingya crisis. While it

[98] Julia Jacobo, "UNICEF warns of scam coronavirus messages circulating through social media", *ABC News*, 10 March 2020 and UNICEF Bangladesh Facebook Page, available at https://www.facebook.com/unicef.bd/photos/a.444413372266529/3852494114791754/?type=3&theater, accessed on 02 April 2020.

[99] See more at Qadaruddin Shishir, "Government Did Not Decide On Opening Schools and Colleges After Eid", Boom, available at https://www.boombd.com/fake-news/govt-did-not-decide-on-opening-schools-and-colleges-after-eid-8991?infinitescroll=1&fbclid=IwAR18NzfIPNns5dpsHEfr3GFJ3ARorsa0SbCIomkZ4RjEwIiLJl0zBoSm0pU, accessed on 22 July 2020.

[100] Associated Press, "Facebook shuts down fake Bangladeshi news sites ahead of vote", *The Daily Star*, 20 December 2018.

[101] Tuhin Shubhra Adhikary and Wasim Bin Habib, "Fake photos trolling", *The Daily Star*, 26 November 2016.

is true that the Rohingyas were facing severe persecution in Myanmar, sharing such kind of content gives rise to information disorder and the severity of the issue is diluted in debates between credible and non-credible information.

4.1.4 *Misleading Content*

Misleading content is a complex and hard-to-define tactic which involves intervention at multiple levels, like the text, image, or statistics. According to First Draft, misleading content is hard to define because it involves several concerns about the context, the amount and significance of the omitted quote, the extent of the photo cropped or edited, and the change in the figures and statistics, etc. This grey area is the reason why in many cases artificial intelligence cannot flag this type of content. This makes such content bypass the filters and dwell in the information ecosystem. This is difficult and time consuming for the fact-checkers as well. It is seen that before the disputed content can be verified, it has already reached a large-scale audience. This is a widely used tactic by many online news portals, click farms, and even cyber troops of political parties. Such content is particularly visible during elections and in times of heated debates between political opponents.

4.1.5 *False Connection*

This is a type of content when headlines, visuals, or captions do not match with the main story. This is again a commonly used tactic by actors like click farms and communication agencies who have incentives to popularize contents very quickly. This is particularly a big problem in South and Southeast Asian regions as a large number of people use Free Basic versions of popular social media platforms, which means that users can have unlimited access to few selected services without using data.[102] So, whenever the users come across a content having a false connection, be it through memes, click baits, and screenshots of headlines, they do not click through it as going to a different website would have data charge. They are also unaware of the option to verify it through search engines as in most cases their understanding of the internet is limited to Facebook or WhatsApp only. So, they fall for the disinformation at first glance without going through the process of reading and analyzing the material and then, forward it as misinformation.

4.1.6 *Satire or Parody*

This kind of content has no intention to cause harm but has the potential to fool. Satire or parody has always been a great tool for artists, activists, and critics but this

[102] Emily Stewart, "Can Facebook be trusted to combat misinformation? Sri Lanka's shutdown suggests no", *Vox*, available at https://www.vox.com/2019/4/23/18511640/facebook-sri-lanka-bombing-social-media-attack, accessed on 08 July 2020.

form of art has been abused in the digital platforms as a tactic of information disorder. Previously in traditional media, people could easily know content to be a satire based on where it is published or who created it. But in online platforms, segments of news or clips of videos are shared from several sources. People cannot trace the origin of the source and fail to see it as a satire. So, in most cases, the repeated share of such content misleads people to think it as legitimate information. Thus, they fall into the misinformation trap. Also, many a time, information from satirical sites either intentionally or mistakenly, is circulated as legitimate information in news portals which further adds to the complexity of the problem.

4.1.7 *Persuasion on Religious Lines*

This is a very effective tactic used by ill-intending actors to disseminate hate, bigotry, and intolerance among religious communities. Such a form of content capitalizes on the sentiment of people. For example, in Bangladesh, messages often go viral across platforms stating that if the receiver is a true believer, s/he needs to forward the message to 100 people (approximately) to give proof of faith, save the religion, etc. Similar tactic is also seen in India over WhatsApp forwards. In this part of the world, such a kind of persuasion largely benefits the ill-intending actors as it drives engagement from a significant section of the users. This practice was dominant in the early days of messaging service and social media use in the countries. Although it has reduced over the years, it still has a large audience during the times of conflicts or tensions among communities.

4.1.8 *Cross-platform Circulation of Contents*

Actors of information disorder usually target to disseminate contents through several platforms. So, the same content is created with modifications based on the features of the platforms, for example, emphasis on narrative on Facebook and WhatsApp, images on Instagram, and audio-visuals on YouTube. Therefore, it is seen that similar messages would circulate through all popular public platforms and messaging services and the audience would repeatedly be exposed to the content. However, the public policy standard and response to disinformation are not the same for all companies. So, even if the content is taken down or fact-checked in a particular platform, in most cases those are still accessible in other platforms. Thus, it makes it difficult to contain the flow of such information.

4.1.9 *Offline Circulation of Online Content*

The connection between online and offline surrounding the spread of disinformation is often overlooked. Countries in South Asia have very strong community-level networks. In many cases in India, Bangladesh and Sri Lanka, it was seen that disinformation campaigns originated online are quickly disseminated through offline

techniques like sharing the printed or photocopied version of the content, announcing through megaphones, mobilizing through mobile phones calls, etc., especially in semi-urban and rural settings. Even if the questionable content is taken down by the intermediaries or regularity organization, the offline version is hard to track. So, it is very important to take the offline circulation under consideration when it comes to disinformation or rumour surrounding critical societal and religious issues.

4.2 Technologies to Amplify Contents

After the content is created, the actors adopt different kinds of tactics to push their messages to their target audience. Sharing of these kinds of content has been possible by using or abusing the technology of digital platforms like search engines, social networking sites, multimedia sharing platforms, blogs, peer to peer messaging applications, etc.

The first tactic that is used is "Astroturfing" which refers to the use of troll factories, click farms, and automated social media accounts to give a false image of grassroots support of the person or thing. It creates inorganic popularity around an individual, organization, or message in the online platforms. Several digital entities are involved in this process.

Firstly, the accounts of cyber troops are used for disseminating the content. Most cyber troops maintain multiple accounts on several platforms so they amplify the messages through those. Next, human-run fake accounts are deployed which conceal the true identity of the user. Fake human-run accounts engage in conversations by posting comments, tweets, or by privately messaging individuals via social media platforms. In the OII's 2019 Global Inventory of Organised Social Media Manipulation, human-operated fake accounts were found in India, Indonesia, Myanmar, Pakistan, the Philippines, Sri Lanka, Thailand, and Vietnam. Fake account networks are usually run in exchange for financial incentives. For example, in the Philippines, community-level fake account operators are paid a fixed daily rate based on a set quota of online posts and comments.[103] Also, previously used spam or marketing accounts are repurposed to spread disinformation. But these activities are usually detected by the social media companies and taken down shortly. So, in order to bypass this, actors are coming up with new and innovative tools. For example, one new tactic is buying legit Facebook accounts from young children and then using it for disinformation purposes.[104] Hacked accounts are also used to spread propaganda and misinformation framing the person. The followers and audiences are usually deceived by the posts or endorsements coming from these accounts. This is commonly seen before the election when accounts of high-profile political candidates are taken over by such activities.

[103] Jonathan Corpus Ong and Jason Vincent A. Cabanes, op. cit.

[104] CNN, "The Fake news machine", available at https://money.cnn.com/interactive/media/the-macedonia-story/, accessed on 03 April 2020.

The information disorder tactic also involves suppressing online voices by taking down content or accounts of targeted individuals by mass reporting. The same tools used for amplifying popularity are deployed to report target content or account through coordinated networks. Posts by activists, political dissidents, or journalists often get reported by such networks of accounts, thus curtailing opposing narratives.

The second tool used in the process of disseminating disinformation is bots. Bots are most commonly used to inflate the number of followers of a social media account. They are also used to amplify narratives or drown out political dissent. In the OII's 2019 Global Inventory of Organised Social Media Manipulation, in South and Southeast Asia, bots activities were reported in Cambodia, India, Indonesia, Malaysia, Myanmar, Pakistan, Philippines, Sri Lanka, and Thailand. Another similar tool is Cybrog which are accounts run in a combination of both humans and bots. While these tools are excessively used in technologically advanced countries like the US, Germany, the United Kingdom, and Russia, in South and Southeast Asia, OII has recorded such activity in Sri Lanka only. However, it is important for other countries to be vigilant.

The third tactic is targeted advertising and boosting of content. By using online and offline sources of data about users, and paying for advertisements on popular social media platforms, some cyber troops target specific communities with disinformation or manipulated media. Here the underlying business models of digital platforms come into play. Just by clicking a few options, actors can set parameters of who they want to target. The platform's algorithms would do the rest. This is a very effective tactic for influencing public opinion before key decision making events like election and referendum.

The fourth tactic is the use of Data analytics. While this tactic came in limelight after the revelation of Cambridge Analytica's operation, actors in different countries are also attempting to use it. Data-driven political consultants play a key role in this regard. While much documentation of this activity in South and Southeast Asia has not been found, Malaysia and India reportedly saw the use of such tactic in recent times.[105] It needs to be noted that some of these tactics are not all illegal or violate the policies of online media platforms as of yet, as the tools used in this process were primarily designed for digital advertisers. But it becomes dangerous when actors of information disorder operate like marketing agencies and abuse the features to disrupt political conversation and influence public opinion through deceptive means.

After the creation and dissemination of disinformation, malinforamtion, and rumour by ill-intending actors, such content is shared by general users of the platforms as misinformation. The combination of features like sharing, reacting, boosting and

[105] See more at Shubhra Pant, "Why this is India's big data election", *Times of India*, 16 April 2019 and Thirteenth Parliament of Singapore, "Report of the Select Committee on Deliberate Online Falsehoods-Causes, Consequences and Countermeasures", available at https://sprs.parl.gov.sg/selectcommittee/selectcommittee/download?id=1&type=subReport, accessed on 05 July 2020.

advertising amplify these contents and make it reach audiences at scales never imagined before. But while abuse of technology is a problem, the underlying technical infrastructure and business models of the online platforms also play a role here.

The Information Age made information from an almost limitless number of sources instantly available at fingertips. But people are not served with such a plethora of information at once. In an online environment, content-shaping and ad-targeting algorithms play the role of human editors and decide the distribution of content. It selects what information and in which order it will be shown to the audience. In this regard, algorithms can be described as certain sets of rules and preferences which determine the order of news feeds or search engine ranks[106], for example, Facebook's News Feed, Twitter's Timeline, and YouTube's recommendation engine. Algorithms are set by social media companies and search engines to determine what is worthy of attention and what should appear in public view.[107] But despite the immeasurable impacts that these content shaping algorithms have in public life, the technology remains largely opaque leaving the public in dark about what shapes their online environment. While Facebook, Google, and Twitter do not hide the fact that they use algorithms to shape content, less is known about how the algorithms actually work, what factors influence them, and how users can customize their own experiences.[108]

Algorithms also determine which users should be shown a given advertisement. According to Open Technology Institute's (OTI) report, the advertiser usually sets the targeting parameters like demographics and presumed interests, but the platform's algorithmic systems pick the specific individuals who will see the ad based on the detailed information that the companies have accumulated about the users and their online behaviour, for example, users' previous interaction with similar content and the interactions of other users who are similar to them.[109] This gives rise to filter bubbles and echo-chambers that actors exploit in disinformation campaigns by initiating targeted contents with the aim to manipulate opinion. Also, while traditional advertising places ads visible to all, like everyone who walks by a billboard or flips through TV channels, see the same advertisement, online targeted advertising can target each audience differently and feed information that actors think would help in influencing their opinion towards specific outcomes. This shapes their information environment and in turn their political reality. Moreover, the business models of most online platforms are optimized for the convenience of advertisers. According to Amnesty International's 2019 report, Google and Facebook's total revenues come almost entirely from advertising, at 84 and

[106] Alan Macleod (ed.), *Propaganda in the Information Age: Still Manufacturing Consent*, London: Routledge, 2019, p. 34.
[107] Ibid.
[108] Nathalie Maréchal and Ellery Roberts Biddle, *It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge*, Washington, DC: Open Technology Institute, 2020.
[109] Ibid.

98 per cent, respectively.[110] This business model seems to facilitate actors of information disorder as the actors operate like any established digital marketing agency.

Here, it can be seen that by following the combination of tactics of content creation and dissemination, actors design different kinds of campaigns to fulfil their objectives. Two factors are crucial in this regard. First, the presence of highly skilled actors capable of deceiving people and at times even the fact-checkers through the content-creation tactics. And second, banking on the underlying technical and business model of the digital platforms to amplify the content and reach audiences, both in mass scale and through micro-targeting. The combination of these creates an information disorder which is increasingly becoming highly complex and difficult to manage.

[110] Amnesty International, *Surveillance Giants: How the Business Model of Google and Facebook Threatens Human Rights*, London: Amnesty International, 2019.

Chapter 5

Impacts of Information Disorder

Information disorder can have impacts on several levels starting from individual to state. In this chapter, the impacts have been analysed from the societal and state level. Here the paper will explore how information disorder can bring harm to communities and impact social cohesion, religious harmony, peace, law and order situation, and political institutions of the countries of South and Southeast Asia.

5.1 Information Disorder Triggering Attacks and Vigilantism

One of the most horrific impacts of information disorder recorded so far has been the 'WhatsApp murders' in India.[111] Between 2017 and 2018, one video of child abduction accompanied by texts indicating that kidnappers are arriving in the city to abduct children went viral over WhatsApp users in India. The video reached some of the remotest parts of the country. The fear-mongering content made people suspicious of travellers and outsiders in their localities. So, whenever an outsider was seen in the locality, many people immediately connected them with the alleged kidnappers and mobilized crowds without verifying or going into the details of the person. Only based on the rumours they received through the messaging services, the angry and suspicious crowd attempted to establish vigilante justice of the alleged child kidnappers. This eventually led to several mob attacks and the murder of at least 30 people in different parts of India.[112] The situation became so volatile that in Tripura, the man hired by the government to tour villages with a megaphone to try to dispel rumours of child kidnappers was also beaten to death. However, in reality, the viral video was the edited version of a child safety and abduction awareness campaign created in Pakistan.[113] But by omitting the context and the end scene in which the boy is returned safely, only the clip of the abduction was shared. The video was widely circulated and it was almost impossible to track down the original creators. This incident eventually led WhatsApp to change its group messaging features for India and also made the government vigilant of such activities, but by this time many lives were already lost due to the impact of information disorder.

Bangladesh also witnessed a similar pattern of information disorder triggered mob attacks in different parts of the country including Dhaka in 2019 which led to the killing of eight people and attack on 30 people.[114] Although the context was different, the fear-mongering surrounding child abduction in online platforms was similar. The

[111] Michael Safi, "'WhatsApp murders': India struggles to combat crimes linked to messaging service", *The Guardian*, 03 July 2018.
[112] Ibid.
[113] Ibid.
[114] "Bangladesh lynchings: Eight killed by mobs over false child abduction rumours", *BBC*, 24 July 2019.

incidents took place after rumours went viral on Facebook, YouTube, and online news portals that human sacrifices were needed to build the Padma Bridge and children were beheaded as offerings. This led to the mobilization of mobs and the killing of people they suspected of kidnapping. One such incident involved a mother of two children who went to a school inquiring about her children's admission process. Suspicion surrounding the rumour people received online, led them to attack and eventually kill her. Later in the statements provided by Bangladesh Police, none of the victims was found to be involved in child abduction.[115] But incidents like this reflected the disastrous impact that information disorder can have in society.

Besides India and Bangladesh, a similar pattern of rumours about kidnappers targeting children to steal their organs became viral in Indonesia. Such rumours also resulted in the loss of lives due to mob attacks on suspected people in many parts of the country between 2010 and 2016.[116] Again in 2018, messages and videos of alleged child kidnappers in action were widely circulated in social media and chat groups, stoking alarm among parents, despite reports confirming that the so-called abduction news was fake.[117] It also needs to be noted that similar incidents of mob killing surrounding a rumour on WhatsApp took place in Mexico in 2018.[118] Mob attacks triggered by similar rumours were also recorded in France in 2019.[119] Here, is it is seen that messaging services, social media platforms, and clickbait journalism of online portals preying on people's fears upgraded the age-old rumour of child abductors to fit for the Information Age by making it spread much faster and with less accountability. In all these cases, the text was associated with video or images of alleged kidnappers fleeing in a speeding car, van, or a motorbike and also disturbing images of dead bodies. But it was later revealed that the videos and images used were from completely separate incidents and placed out of context. But with the speed in which such contents are shared and due to the fear-mongering narrative of the messages, most people easily fell for it without checking its authenticity. A section of them was triggered by these contents to deliver vigilante justice, often resulting in deadly consequences.

Information disorder also triggered mob attacks and targeted killings in Pakistan. Rumours, misinformation and disinformation surrounding polio vaccination are widely persistent in Pakistan, which is one of the only three countries where polio remains endemic.[120] Propaganda and rumours against polio vaccination, for example,

---

[115] Ibid

[116] Nils Bubandt, "From Head-hunter to Organ-thief: Verisimilitude, Doubt, and Plausible Worlds in Indonesia and Beyond", *Oceania*, Volume 87, Issue 1, 2017.

[117] A. Muh. Ibnu Aqil, "Fake abduction videos stoke fear among parents", *The Jakarta Post*, available at https://www.thejakartapost.com/news/2018/11/01/fake-abduction-videos-stoke-fear-among-parents.html, accessed on 01 July 2020.

[118] Marcos Martínez, "Burned to death because of a rumour on WhatsApp", *BBC*, 12 November 2018.

[119] Aurelien Breeden, "Child Abduction Rumors Lead to Violence Against Roma in France", *The New York Times*, 28 March 2019.

[120] The Global Polio Eradication Initiative, "Endemic Countries", available at http://polioeradication.org/where-we-work/polio-endemic-countries/, accessed on 13 June 2020.

children falling sick after being vaccinated, oral vaccine constituted of hormones to make children sterile, the product was *haram*[121], it was a cover for foreign interests, etc. were widely circulated.[122] These resulted in fatal consequences in parts of the country, particularly in the western regions bordering Afghanistan. Such rumours not only risked several thousand children getting affected but, it also endangered the lives of polio vaccinators and the law enforcement agency officials deployed to protect the vaccinators. In the past few years, rumours and disinformation campaigns widely spread through social media triggered mob attacks and target killings of several on duty polio workers and police officers guarding vaccinators.[123]

It was seen that in many cases the disputed anti-vaccination content was downloaded from Facebook and then circulated through either WhatsApp or directly transferred from one cell phone to another,[124] thus even when the contents were taken down from Facebook, it was still available in other media. In April 2019, several staged videos were posted to Twitter and Facebook from the province of Khyber Pakhtunkhwa which falsely reported children falling gravely ill after receiving the polio vaccine.[125] According to First Draft, the videos were part of an anti-vaccine disinformation campaign that caused mass panic in the country. In the same week, as the videos spread, a mob of 500 people set fire to a health clinic in Peshawar and mosque loudspeakers broadcast rumours that polio medicines were 'poisonous'.[126] Here, both online and offline tactics were used to disseminate the messages. Besides threatening the lives of the individuals involved, such campaigns seemed to have an impact on the already difficult battle against the virus. In 2018, only 12 polio cases were reported in Pakistan but in 2019, the number spiked to 146 cases with 92 in the province of Khyber Pakhtunkhwa, where the videos originated.[127]

Rumours surrounding sterilization also triggered mob attacks in Sri Lanka. In early 2018, a false claim that 23,000 sterilization pills were seized from a Muslim pharmacist in Ampara, Sri Lanka was spread through social media platforms.[128] This was seen by many Sinhalese as a Muslim plot to sterilize and destroy the Sinhalese population. While this rumour spread online, it also had real life implications on the very next day.

---

[121] Forbidden by Islamic Law.

[122] Lucy Lamble, "Killings of police and polio workers halt Pakistan vaccine drive", *The Guardian*, 30 April 2019.

[123] See more at Ben Farmer, "Motorbike gunmen kill two polio workers in Pakistan", *The Telegraph*, 30 January 2020 and Haroon Janjua, "Polio vaccine worker shot dead is third killed in Pakistan this week", *The Telegraph*, 26 April 2019.

[124] Haroon Janjua, "Pakistan thanks Facebook for prompt removal of anti-vaccine posts", *The Telegraph*, 26 June 2019.

[125] Lydia Morrish, "How fake videos unravelled Pakistan's war on polio", First Draft, available at https://firstdraftnews.org/latest/how-fake-videos-unravelled-pakistans-war-on-polio/, accessed on 13 June 2020.

[126] Ibid.

[127] Pakistan Polio Eradication Programme, "WPV Polio Cases Across Pakistan's Provinces", available at https://www.endpolio.com.pk/polioin-pakistan/polio-cases-in-provinces, accessed on 13 June 2020.

[128] Amanda Taub and Max Fisher, op. cit.

Sinhalese customers dining in a Muslim run restaurant found a suspicious object on food and immediately connected it to the sterilization rumour. The event quickly escalated and led to the mobilization of mobs who attacked the owner, destroyed the restaurant and torched a local mosque.[129] The brawl over the food was also recorded and shared online, further amplifying tension between the Muslims and Buddhists.

Another incident involving sterilization rumour occurred in Sri Lanka in May 2019 where a Muslim doctor was alleged for secretly sterilizing 4,000 Sinhala Buddhist women.[130] First published in a Sinhala language newspaper, the story became widely circulated through social media. Later, the claim was debunked as false after arrest and investigation of the alleged doctor by Sri Lanka authority.[131] But such rumours have serious implications in society. According to Reuters report, "Allegations a Muslim doctor might be forcibly sterilizing Buddhists are particularly incendiary on an island where hardliners within the Buddhist majority have accused Muslims of seeking to use a higher birth rate to spread their influence".[132] In both the cases discussed here, it can be seen how a rumour can take a communal spin, create tension among communities, and also trigger mob attacks against people of a target community.

These incidents from India, Bangladesh, Sri Lanka, Pakistan, and Indonesia show the magnitude of the impact that a rumour on social media can result in. Here it is seen that the narrative and content may be different, but in all cases, sections of general people decided to take the matter into their hands and deliver justice instead of relying on the state institutions. This is a critical problem for this region. Thus, it is very important to closely analyze the socio-political, cultural, and religious context of online media rumours to assess which kind of contents have the potentials to trigger attacks and vigilantism and therefore, take necessary precautionary measures.

5.2 Hate Speech Inciting Violence Against Selected Communities

Hate speech inciting attacks against a community or escalating on-going conflicts is not a new problem, but using the strength of the online platforms to trigger, amplify and accelerate the process is a major problem of information disorder. Such actions are widely seen in many parts of the world, especially targeting vulnerable or minority communities. In South and Southeast Asian regions, there is evidence of hate speeches on social media platforms to be the causal factors of violence against target communities, in most cases the ethnic and religious minority groups in the country.

---

[129] Ibid.

[130] AFP Sri Lanka, "Sri Lankan authorities found the Muslim surgeon had not performed any sterilisations", *AFP Fact Check*, 05 July 2019.

[131] Ibid.

[132] Alexandra Ulmer and Omar Rajarathnam, "Unsubstantiated claims Muslim doctor sterilized women raise tensions in Sri Lanka", *Reuters*, 06 June 2019.

Myanmar is an important case where social media was used to incite violence and hatred against the Rohingya minority group. Marzuki Darusman, Chairman of the United Nations Independent International Fact-Finding Mission on Myanmar stated that social media had played a "determining role in Myanmar and ... substantively contributed to the level of acrimony and dissension and conflict, if you will, within the public. Hate speech is certainly of course a part of that."[133] The Human Rights Council's report states, "The role of social media is significant. Facebook has been a useful instrument for those seeking to spread hate, in a context where, for most users, Facebook is the Internet. Although improved in recent months, the response of Facebook has been slow and ineffective. The extent to which Facebook posts and messages have led to real-world discrimination and violence must be independently and thoroughly examined."[134]

The series of incidents surrounding violence and oppression against the Rohingya community led to around one million people fleeing the country and taking shelter in bordering Bangladesh. It is widely believed that the combination of misinformation, disinformation, malinformation, and hate speech by influential actors in the country contributed to violence against the community in this prolonged period of conflict. Multiple actors played a role in this process, like the military, radical Buddhist groups, and religious zealots. To scrutinize the role of the military, a request was filed to Facebook asking for communication materials of Myanmar military officials and police forces as part of the ongoing International Court of Justice case.[135] Hard-line Buddhist monks like Ashin Wirathu are also known to be responsible for disseminating hate speech, inflammatory content and sowing divisions in the society, as discussed in Chapter 3. A section of general citizens influenced by these hard-line Buddhist groups also resonated with the messages and amplified it through their accounts. The online activities of all these actors are known to have played a critical role in inciting attacks and escalating the ongoing conflict against Rohingya communities.

Following the atrocities and under severe criticism from civil society and international organizations, Facebook took steps to ban some of the prominent actors of information disorder in the country and combat hate speech. However, an investigation of Reuters found that many of these networks were still active and propagated hate against Rohingya and other Muslims in Myanmar. As of August 2018, the investigation found more than 1,000 examples of posts, comments, images and videos attacking the community.[136] Additionally, it was found that, unlike the previous time, Facebook was not the only platform used for this purpose. The disinformation and hate campaign also proliferated Twitter. This shows the widespread information

[133] Tom Miles, "U.N. investigators cite Facebook role in Myanmar crisis", *Reuters*, 13 March 2018.
[134] Human Rights Council, *Report of the Independent International Fact-Finding Mission on Myanmar*, Geneva: Human Rights Council, 2018, p. 14.
[135] "U.S. court asked to force Facebook to release Myanmar officials' data for genocide case", *Reuters*, 10 June 2020.
[136] Steve Stecklow, *Inside Facebook's Myanmar operation: Hatebook*, Yangon: Reuters, 2018.

disorder surrounding the Rohingyas in Myanmar and the magnitude of the impact it can have in real life.

A similar pattern of hate speech inciting violence was also recorded in Sri Lanka. In Sri Lanka, the government accused Facebook of failing to control rampant hate speech that it says contributed to anti-Muslim riots that left three people dead and the country under a state of emergency.[137] The violence in Kandy is understood to have been sparked when a group of Muslim men in Digana town was accused of killing a man belonging to the majority Sinhala Buddhist community and in response, violence erupted targeting the Muslim community.[138] In the whole process, social media played a key role. During the event, rumours and hate speech spread like wildfire on Facebook and perpetrators carried out arson attacks, targeting dozens of mosques, shops and homes of Muslims. The situation escalated to a level where the government had to briefly ban Facebook.[139] Amith Weerasinghe, the leader of the Buddhist hard-line group Mahason Balakaya was accused of helping to instigate the violence through nearly 150,000 followers on his Facebook page.[140]

In response to the events in Sri Lanka, Facebook hired Article One, a human rights consultancy firm to investigate the matter. Article One report revealed, "Facebook platform contributed to spreading rumours and hate speech, which may have led to 'offline' violence. Indeed, the assessment found that the proliferation of hate speech (e.g., "Kill all Muslims, don't even save an infant; they are dogs") and misinformation (e.g., that a Muslim restaurateur was adding sterilization pills to his customers' food) may have contributed to unrest and in the case of the restaurateur and others, physical harm."[141]After the findings were released, Facebook acknowledged its impact in Sri Lanka stating, "We recognise, and apologise for, the very real human rights impacts that resulted."[142]

While these two cases in Myanmar and Sri Lanka received national and international attention due to the magnitude of the impact, hate speech against target communities remains widely persistent on social media platforms across many countries in the regions. Several times it is seen that such content continued to stay in the platforms despite being reported by users and fact-checkers. This aspect of information disorder needs more attention from all relevant stakeholders like social

[137] Michael Safi, "Sri Lanka accuses Facebook over hate speech after deadly riots", *The Guardian*, 14 March 2018.
[138] Michael Safi and Amantha Perera, "Sri Lanka declares state of emergency after communal violence", *The Guardian*, 06 March 2018.
[139] Meera Srinivasan, "Online hate and its offline costs", *The Hindu*, 16 May 2020.
[140] Michael Safi, op. cit.
[141] Article One, "Assessing the Human Rights Impact of the Facebook Platform in Sri Lanka" available at https://about.fb.com/wp-content/uploads/2020/05/Sri-Lanka-HRIA-Executive-Summary-v82.pdf, accessed on 22 July 2020.
[142] Joshua Brustein, "Facebook Apologizes for Role in Sri Lankan Violence", *Bloomberg*, 13 May 2020.

media companies, fact-checkers, civil society organizations and associated government agencies to prevent online contents from creating offline violence.

5.3 Religious Defamation Triggering Revenge Attacks

Another pattern of information disorder which can instigate an attack on communities is religious defamation. In Bangladesh, religious defamation and rumours about defamation being shared online triggered conflict and tension between the majority Muslims and the minority communities on several occasions in the past ten years. Here a combination of both online and offline tactics was used in the process. The atrocity in Ramu, an Upazila in the southern part of Bangladesh at Ukhia, Cox's Bazaar was one of the worst acts of communal attacks in Bangladesh. It was one of the first noticeable incidents of instigating hate campaigns through social media. On 29 September 2012, as many as 18 pagodas were destroyed and about 50 houses burnt down in Ramu and nearby areas by the mobs who were Muslims.[143] A fabricated photo hurting the religious sentiment of Muslims allegedly shared by a Buddhist man was considered to be the reason behind this orchestrated attack on the minority Buddhist community.[144] A similar incident occurred in Pabna in 2013. A Hindu student named Rajib Shah was accused of maligning the Prophet and on that basis, atrocity was carried out on the predominately Hindu village in Santhia upazila of the district.[145] The attack was incited when photocopies of a Facebook page defaming the Prophet were circulated in the village linking it to that student Rajib Shah. However, The Daily Star analyzed the disputed content and could not find any direct connection with the accused student.[146]

On 30 October 2016, Nasirnagar, the district in Brahmanbaria, witnessed another incident where violence was triggered by a Facebook post demeaning the Holy Kaaba purportedly from the account named Rasraj Das.[147] However, investigation found that the photo was not uploaded from Rasraj's phone, instead, it might have been used to frame him.[148] One year later, on 10 November 2017, violence erupted in Thakurpara village of Rangpur when rumour spread that Titu Roy, a Hindu man of the village, defamed the Prophet in a Facebook post. Based on the rumour, at least 30 Hindu houses were looted, vandalized and torched by zealots.[149] During the incident, one man was killed and 20 others were injured as police fired rubber bullets and teargas

[143] Julfikar Ali Manik, "A devil's design", *The Daily Star*, 14 October 2012.
[144] Inam Ahmed and Shakhawat Liton, "Ramu violence: In the shadow of what we don't know", *The Business Standard*, available at https://tbsnews.net/bangladesh/crime/ramu-violence-shadow-what-we-dont-know, accessed on 11 August 2020.
[145] Ahmed Humayun Kabir Topu, "Hindus attacked in Pabna", *The Daily Star*, 3 November 2013.
[146] Ibid.
[147] "Mayhem in B'baria", *The Daily Star*, 31 October 2016.
[148] Shakhawat Liton, "Rasraj a victim", *The Daily Star*, 04 December 2016.
[149] "B'baria-style plot behind Rangpur arson", *The Daily Star*, 15 November 2017.

shells to restore law and order.[150] But in an investigative report of Dhaka Tribune, it was found that the post was uploaded from an account named Md Titu, not Titu Roy. It is likely that someone else might have impersonated him.[151]

The most recent incident along a similar narrative was recorded in Bhola's Borhanuddin Upazila on 20 October 2019. Clashes broke surrounding screenshots of an alleged Hindu youth engaging in defamatory conversation against Islam on Facebook messenger. The content became viral in the locality and led to the mobilization of hardliner local Muslims under the banner "*Sarbastorer Muslim Tawhidi Janata*" to stage demonstrations demanding trial of the youth. As an act of revenge for the defamation, a house was torched and 12 more vandalized belonging to the Hindu community.[152] Clashes also broke out between the law enforcement agency deployed to control the situation and the demonstrators which eventually resulted in the death of four people and injured more than a hundred. In this case too, it was found that the Facebook account of the alleged Hindu youth was hacked and he was allegedly framed for religious defamation.[153]

In all these cases it is seen that the allegation of defamation was based on fabricated content, hacked account or Facebook pages used to frame a person. Although the defamation originated from activities or allegation of activities online, it is seen that the conflict was triggered primarily through offline measures like announcements, sharing of photocopies of content, pamphlets or screenshots. In all these cases, different minority communities became victims under different circumstances, but it was all centred on online activities that hurt the religious sentiments of the majority Muslim population. This pattern of online religious defamation triggering revengeful actions needs to be carefully studied as it can have severe consequences in societies. It also creates the scope of framing people as it was seen in most cases that the alleged person did not have a direct connection with the disputed content, rather the person was a victim of the situation.

5.4 Disinformation Fuelling Xenophobia

In multicultural and heterogeneous countries of Southeast Asia, online platforms are often used to fuel xenophobia. Such acts are increasingly becoming a concern in countries which hosts a large number of migrants, like Singapore and Malaysia.

The Real Singapore (TRS), a socio-political website in Singapore, is one such online platform which was responsible for creating several anti-foreigner content and

[150] Ibid.

[151] Liakat Ali Badal and Kamrul Hasan, "Rangpur Attack: FB post uploaded from Rangpur, Titu lives in N'gang" *Dhaka Tribune*, 12 November 2017.

[152] "Police-Protesters Clash in Bhola Over Hate Spread Thru' FB ID: 4 Killed, 100 Injured", *The Daily Star*, 21 October 2019.

[153] Arifur Rahman Rabbi and Ahad Chowdhury Tuhin, "Bhola clash: Hacker of Biplob's Facebook account identified, Kamal says", *Dhaka Tribune*, 24 October 2019.

encouraging hostility against them by propagating falsehoods since 2012.[154] This fuelled xenophobia on one hand and brought in significant revenue from online advertisement to the website owners on the other. In 2016, TRS' founders were found guilty of sedition and deliberately sowing discord between Singaporeans and foreigners.[155]

Disinformation to instigate xenophobia was also seen in Malaysia. For example, during the 2013 general elections in Malaysia, disinformation spread that "40,000 Bangladeshi nationals were brought to Malaysia to vote to help swing the votes to the benefit of the then ruling coalition".[156] Later, the Malaysian Communications and Multimedia Commission (MCMC) identified this as false news and took the initiative to arrest the suspect behind this disinformation campaign.[157] However, the information had already been widely circulated through social media. The impact of that disinformation was felt by those whose physical appearance looked like foreigners as they were confronted and manhandled in polling stations.[158]

5.5 Impacts of Information Disorder on Social Movements and Protests

The Information Age brought significant changes in how social movements are initiated and carried on. Using data from the Global Database of Events, Language, and Tone, the Center of Strategic and International Studies (CSIS) published a report which revealed that mass protests increased annually by an average of 11.5 per cent from 2009 to 2019 across all regions of the world.[159] This rise in number seemed to have a connection with the penetration of the internet. In 2009, 1.5 billion people were connected to the internet, in 2019, it became more than double and the number of internet users reached 4 billion. While digital connectivity cannot be said to be the sole driving factor of the overall trend in protest, however, the internet connectivity facilitating rapid transmission of information became a critical enabler for global protests. Social media and virtual discussion boards served as platforms for sharing grievances, connecting aggrieved people, and ultimately spurring mass mobilization. But at the same time, digital media also became the reason why many social movements boiled down to chaotic information disorders and created public unrest.

---

[154] Norman Vasu, Benjamin Ang, Terri-Anne-Teo, Shashi Jayakumar, Muhammad Faizal, and Juhi Ahuja, *Fake News: National Security in the Post-Truth Era*, Singapore: S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University, 2018.

[155] Ibid.

[156] Moonyati Mohd Yatid, "Truth Tampering through Social Media: Malaysia's Approach in Fighting Disinformation & Misinformation", *The Indonesian Journal of Southeast Asian Studies*, Volume 2, No. 2, 2019.

[157] Beatrice Nita Jay, "Culprit Who Viralled False News of Bangladeshi Phantom Voters with Blue Caps on to be Arrested", available at https://www.nst.com.my/news/nation/2018/05/365039/culprit-who-viralled-false-news-bangladeshi-phantom-voters-blue-caps-be, accessed on 15 July 2020.

[158] Moonyati Mohd Yatid, op. cit.

[159] Samuel J. Brannen, Christian S. Haig, Katherine Schmidt, *The Age of Mass Protests: Understanding an Escalating Global Trend*, Washington D.C.: Center For Strategic & International Studies, 2020.

In Bangladesh, the impact of information disorder was visible in all major movements and protests in the past few years. The 2013 Shahbagh movement was one of the first large-scale movements organized through social media. While the Shahbagh movement saw several social media platforms being positively used to organize the movement and put forward demands related to the verdict of war criminals of Bangladesh's liberation war, it also received a counter-attack from anti-liberation forces and extremist groups. These groups carried out disinformation campaigns on social media to disrupt the support for the movement. Several contents like manipulated news, fake captions, and doctored images were spread through social media to disrupt the protest.

The two major events of 2018, the Quota Reform Movement in April and the Road Safety Movement in August were also crippled by information disorder circulating in social media. Social media sites like Facebook were initially used as a resource for these movements, but it gradually turned into a battleground between the keyboard warriors of the groups in favour and against the movement. So, the digital platforms that had an immense contribution to organize these movements, were also the ones to disrupt it. Some of the tactics of disinformation involved sharing of unverified news through Facebook group posts and status updates to intensify the protest, spreading rumours to disrupt the movement, reporting and bringing down account and pages of activists, disinformation to defame political organizations, reusing old photos to support false news, producing distorted videos, misinterpretation of news and using Facebook Live videos to spread panic. This caused fear, chaos, and confusion among the protestors and also impacted the law and order situation in the country. General people who were not part of the protest also contributed to the information disorder through sharing misinformation. The combination of all these agitated the citizens and triggered conflicts in different parts of the country resulting in days of unrest.

In India, protests against the Citizenship Amendment Act starting in 2019 also triggered extensive distribution of misinformation and disinformation across different social media platforms and messaging services.[160] Just in the first few weeks of protests, one social media platform reportedly flagged and removed around 2,500 fake news and communal items.[161] Contents like these have potential risks for further escalating the on-going tension and unrest. Thus, such contents need to be carefully scrutinized but at the same time, ensuring that the legitimate voices of protestors and activists are not curtailed.

5.6 Ramifications for Democracy

Information disorder in the political sphere has been one of the most widely discussed and debated topics at present. Joseph S. Nye Jr. referred that politics in an

[160] Anumeha Chaturvedi, "2019—The year of fake news", *The Economic Times*, 20 December 2019.

[161] Karan Choudhury and Neha Alawadhi, "CAA protests: 15,000 social media mediators fight to root out fake news", *Business Standard*, available at https://www.business-standard.com/article/companies/caa-protests-15-000-social-media-mediators-fight-to-root-out-fake-news-119121601331_1.html, accessed on 15 July 2020.

Information Age is ultimately about whose story wins.[162] Besides international politics, this trend is also visible in domestic politics as political parties compete to win over hearts and minds and create a favourable image among the public, particularly during elections. But in doing that, many a time, political parties, politicians, cyber troops, and trolls have resorted to the use of disinformation. This has started to affect the core elements of democracy and governance. On the one hand, the abundance of accessible information allowed the government and politicians to be open, transparent and interactive and, empowered citizens both individually and collectively to shape the institutions whose decisions impact their lives.[163] But on the other, it has also created an environment where actors starting from foreign states to individual citizens can initiate disinformation campaigns to secure certain strategic and political objectives. Social media is one of the most influential platforms in this regard. Philip Howard, Director at the OII, states that social media has made democracy weak.[164] Social media platforms seem to have challenged the information ecosystem of several democracies as seen in different incidents over the past few years.

Ironically, the very democratic principles that ensure free and fair speech is compromised when the same principles allow for the spread of disinformation.[165] While initially the fact that digital platforms could be used to manipulate decision making in votes was discarded by the social media companies, gradually they admitted the power of these platforms in challenging democracy. In discussing what effect social media have on democracy, Facebook's Product Manager of Civic Engagement, Samidh Chakrabarti states, "If there's one fundamental truth about social media's impact on democracy it's that it amplifies human intent—both good and bad. At its best, it allows us to express ourselves and take action. At its worst, it allows people to spread misinformation and corrode democracy."[166] While the discussions regarding disinformation campaigns surrounding the 2016 US elections and Brexit referendum have dominated the discussions, the Global South has also been deeply affected by such actions. While this is a problem for all political systems in this region, it has been particularly visible in the democracies of South and Southeast Asia. The actors and tactics might be different, but the objective to manipulate public opinion, sow discord, and reinforce pre-existing bias against communities through digital tools is similar. The following is a brief overview of information disorder witnessed by some of the countries during recent elections.

---

[162] John Arquila and David Ronfeldt, *The Emergence of Noopolitik: Toward an American Information Strategy*, Santa Monica: RAND, 1999 in Joseph S. Nye, Jr, *Soft Power: The Means to Success in World Politics,* New York: Public Affairs, 2004.

[163] Laura Chinchilla, "Post-Truth Politics Afflicts the Global South, Too", *The New York Times*, 18 October 2019.

[164] "Is Social Media Killing Democracy? Computational Propaganda, Algorithms, Automation and Public Life", available at https://www.oxfordmartin.ox.ac.uk/videos/is-social-media-killing-democracy-computational-propaganda-algorithms-automation-and-public-life/, accessed on 04 June 2020.

[165] Juhi Ahuja, "Fake News and India's Democracy", *The Diplomat*, 02 June 2018.

[166] Samidh Chakrabarti, "Hard Questions: What Effect Does Social Media Have on Democracy?", Facebook Newsroom, available at https://newsroom.fb.com/news/2018/01/effect-social-media-democracy/, accessed on 04 June 2020.

India experienced various forms of information disorder in the past few elections. With a series of disinformation, misinformation, and malinformation from several actors circulating through different platforms, technology was abused to exploit and amplify existing fault lines like religious nationalism, caste politics, and political polarization. The fact-checking organization Alt News has estimated that during the 2019 elections, the spread of misinformation increased by 40 per cent compared with non-election times.[167] Information disorder through images was also widely disseminated besides text messages. Two Massachusetts Institute of Technology (MIT) researchers studied messages of several politically-oriented WhatsApp groups prior to the 2019 Indian national election to find the level of image-based misinformation shared through those. The study involved 2,000 most shared images within the groups and 500 randomly selected images and found that 13 per cent of the images shared were misinformation.[168] The images could be broadly divided into three categories, i.e., old images taken out of context, memes, and photoshopped images. Qualitative analysis of the images revealed that those mostly covered topics like historic and religious myths, nationalism, and political memes. The study also found that many images conveyed a sense of urgency and shock value making them spread faster.

In the Philippines, the disinformation campaign was rampant in the 2016 elections where the political candidates deployed cyber troops and used disinformation as one of the key tactics. In discussing this new trend, Yvonne T. Chua and Ma. Diosa Labiste from the University of the Philippines opined that "The 2016 Philippine presidential election brought to the fore how information disorder can transcend traditional platforms to permeate the internet, especially social media, and influence public opinion through tampered or manufactured reality".[169]

In Indonesia, information disorder along political lines was also widespread across online media platforms, particularly on Facebook and Twitter. Both politically and financially motivated actors like social media consultants and cyber troops, also known as 'buzzers' were widely operational during elections.[170] These actors have reportedly used disinformation playing on ethnic and religious sentiments to undermine election candidates.[171]

The information disorder has also resulted in declining trust in the government, electoral system, and media. Disinformation tactics have often led to convincing people

[167] Aria Thaker, "India's Fake-News Crisis has Intensified During the 2019 Elections, Say Fact-Checkers", available at https://qz.com/india/1609763/alt-news-boom-live-on-fake-news-detection-amid-indian-election/, accessed on 11 August 2020.
[168] Kiran Garimella and Dean Eckles, "Images and Misinformation in Political Groups: Evidence from WhatsApp in India", *The Harvard Kennedy School (HKS) Misinformation Review*, 2020.
[169] Yvonne T. Chua and Ma. Diosa Labiste, op. cit.
[170] Fanny Potkin and Agustinus Beo Da Costa, "In Indonesia, Facebook and Twitter are 'buzzer' battlegrounds as elections loom," *Reuters*, 13 March 2019.
[171] Thirteenth Parliament of Singapore, op. cit.

that real source of information coming from the government or established media cannot be trusted, thus further complicating the information ecosystem.

5.7 Impacts on Foreign Relations

Misinformation and disinformation shared against a foreign country, its policies, and leaders can affect bilateral relations between the countries. This was recently visible in incidents surrounding Bangladesh, Bhutan, India, and Nepal. Following the boundary issue between Nepal and India, misinformation and disinformation were spread by social media users claiming of border skirmishes between the two countries and shooting down of jets.[172] The claim was accompanied by images with false connections. However, the fact-checking organization Alt News debunked such claims and published a detailed report on it.[173] But such kind of false information going viral in a time when the two countries are going through slight complexities, have the potential to impact bilateral relations. A similar incident was also reported surrounding the issue of shared water bodies with Bhutan. Posts were circulated among social media users claiming that Bhutan has stopped the supply of irrigation water to farmers in Assam.[174] The news also included misleading photos. In response, Bhutan's Ministry of Foreign Affairs issued a press release clarifying its position and pointed out that such misinformation can cause misunderstanding between the people of the two countries.[175] India Today's Anti Fake News War Room also found the information to be misleading and published a report on it.[176]

In recent times, misinformation was also shared regarding foreign leaders. Such news has the potential to harm bilateral relations and negatively impact the expatriate community. For example, in Bangladesh, misinformation was spread in online portals and social media wrongly quoting the Italian Prime Minister labelling the Bangladeshis as 'virus bomb'.[177] In response, the Ministry of Foreign Affairs in Bangladesh issued a press release clarifying the misquotation and urged the media to check the authenticity of news before sharing. The fact-checking organization Boom also published a detailed report debunking the allegation on the Italian Prime Minister.[178] Although the incidents

[172] Archit Mehta, "Old images falsely shared as Indian fighter jets shot down by Nepal", Alt News, 25 July 2020, available at https://www.altnews.in/old-images-falsely-shared-as-indian-fighter-jets-shot-down-by-nepal/, accessed on 11 August 2020.

[173] Ibid.

[174] Chayan Kundu, "Fact Check: Has Bhutan deliberately blocked irrigation water to Indian farmers in Assam?", *India Today*, available at https://www.indiatoday.in/fact-check/story/has-bhutan-deliberately-blocked-irrigation-water-to-indian-farmers-in-assam-1693016-2020-06-26, accessed on 11 August 2020.

[175] Ministry of Foreign Affairs, *Clarifications on the Recent News Articles Published in India Alleging that Bhutan has Stopped the Supply of Irrigation Water to Farmers in Areas in Assam Adjoining Samdrup Jongkhar District*, Thimphu: Ministry of Foreign Affairs, Royal Government of Bhutan, 2020.

[176] Chayan Kundu, op. cit.

[177] Ministry of Foreign Affairs, *Press Release: Some Newspapers Misquoted Italian Prime Minister*, Dhaka: Ministry of Foreign Affairs, Government of Bangladesh, 2020.

[178] Qadaruddin Shishir, "Italy's Giuseppe Conte Did Not Say Each Bangladeshis Are A Viral Bomb", Boom, 11 July 2020, available at https://www.boombd.com/fake-news/italys-giuseppe-conte-did-not-say-each-

discussed here were promptly addressed by the authorities, it is still important for the governments to be vigilant of such kind of information on online platforms in the future and take proactive measures.

5.8 Impacts of the COVID-19 'Infodemic'

Information disorder has taken a new dimension following the outbreak of the coronavirus disease across the world in late-December 2019. While the World Health Organization (WHO) and relevant bodies battle to address this pandemic in the real-life, they face an equally challenging battle in the virtual world as disinformation campaigners, conspiracy theorists, and opportunists flood the digital media with misinformation, disinformation, malinformation, and hate speech. Starting from debates regarding who created the virus to whether or not to accept medical support, the digital world witnesses a concerning impact of the information disorder. The WHO labelled this over-abundance of information—some accurate and some not, as a massive 'infodemic' which makes it hard for people to find trustworthy sources and reliable guidance when they need it.[179] Online platforms are central to this information conundrum. With most of the people staying home have limited circulation of trusted media like newspapers, either due to lock down or voluntary decline of subscription due to fear of contamination, a large percentage of the people resorted to online media platforms as their primary source of information. Several forms of information disorder dominated their social media newsfeed in the early days of the crisis resulting in an 'infodemic'.

The COVID-19 infodemic has multiple dimensions. First, is the disinformation surrounding the origin of the disease. The digital media is flooded with parallel narratives where one side believes that the coronavirus had been created in a secret government lab in China[180] while another narrative alleges that the virus is a biological weapon manufactured by the Central Intelligence Agency (CIA).[181] Alternatively, there were few more narratives regarding the origin of the virus including the one that promotes that coronavirus was an invention of the pharmaceutical industry, intended to sell expensive drugs and more vaccines to the public.[182]

While many aspects of this never-ending battle regarding the origin of the virus can be brushed off as a hoax, there are few aspects of this debate which have serious implications on people's lives. Disinformation campaigns like this make it harder to

bangladeshis-are-a-viral-bomb-8832?infinitescroll=1, accessed on 11 August 2020.

[179] World Health Organization, *WHO Novel Coronavirus(2019-nCoV) Situation Report—13*, Geneva: World Health Organization, 02 February 2020.

[180] Sheera Frenkel, Davey Alba and Raymond Zhong, "Surge of Virus Misinformation Stumps Facebook and Twitter", *The New York Times,* 08 March 2020.

[181] Jessica Glenza and agencies, "Coronavirus: US says Russia behind disinformation campaign", *The Guardian*, 18 March 2020.

[182] Sheera Frenkel et al., op. cit.

respond to the crisis. Spreading multiple and often contradictory versions of events undermine trust in objective facts and give rise to coronavirus deniers. Also, there are several forms of disinformation and misinformation which discourages patients from receiving treatment. This is a major concern in the efforts to contain this disease. On the flip side, thousands of contents are also emerging regarding preventive measures and treatment of this disease. While many of the information available might seem harmless, many posts promoting immunity boosting drinks turned out to be dangerous and life-threatening. The Food and Drug Administration (FDA) referred to one 'miracle mineral solution' posted many times on Facebook and Twitter as 'the same as drinking bleach.'[183]

In many countries of South Asia, such infodemic was shared with a communal spin. Several posts were seen on social media which claimed that this virus would not affect certain religious communities, thus preventing them from taking precautionary measures. Few groups of religious leaders and their followers have often resorted to misinformation and disinformation surrounding the topic to popularize their sermons. This was particularly evident in the early days of the COVID-19 pandemic. At least 60 such religious leaders came up with various unscientific comments in Bangladesh.[184] Their messages varied from fear-mongering, spreading panic and intolerance, advocating for unverified treatments, and assuring that Muslims were immune to this virus. As these preachers are blindly followed by a large number of people, distorted information from them can have serious implications on health and religious harmony. Alternately, the Muslims in India and Sri Lanka have been stigmatized surrounding the same issue by groups of Hindu and Buddhist hardliners respectively.[185] They blamed the minority Muslim communities of the countries for spreading the virus and disseminated several unverified allegations through different media platforms. A study by Equality Labs on the Islamophobic COVID-19 hate speech in South Asia found that hate speech and disinformation targeting Muslims runs rampant across Twitter, Facebook, WhatsApp, and other social media platforms. The study reported that in India, such content grew from established Islamophobic social media accounts, pages and groups of Hindu nationalists, and had common themes such as Muslims depicted as the virus; Muslims equated with bioterrorism, with the weapon being the virus; False claims that Muslims are intentionally spreading the disease to non-Muslims and Hindus as a form of 'jihad', etc.[186]

---

[183] Ibid.

[184] Zia Chowdhury, "No steps yet to stop 'Islamic scholars' from spreading disinformation on Covid-19", *The Business Standard*, available at https://tbsnews.net/coronavirus-chronicle/covid-19-bangladesh/no-steps-yet-stop-islamic-scholars-spreading?fbclid=IwAR1AWP3VZ3eGjSUarml1wwDwNK6mHgh3-kO36OnLCgwPTm0bsUksFx2R91o#.Xr-cl1XFBAQ.facebook, accessed on 02 June 2020.

[185] Omar Suleiman, "Like India, Sri Lanka is using Coronavirus to Stigmatise Muslims", *Al Jazeera*, available at https://www.aljazeera.com/indepth/opinion/india-sri-lanka-coronavirus-stigmatise-muslims-200519134939934.html, accessed on 02 June 2020.

[186] T. Soundararajan, A. Kumar, P. Nair and J. Greely, "Coronajihad: An Analysis of Islamophobic COVID-19 Hatespeech", Equality Labs, 2020, available at https://www.equalitylabs.org/coronajihad, accessed on 18 July 2020 and Billy Perrigo, "It Was Already Dangerous to Be Muslim in India. Then Came the Coronavirus", *Time*, 03 April 2020.

To tackle this infodemic, social media companies in collaboration with WHO and relevant organizations took visible initiatives to make the correct information available to people starting from mid-March 2020. However, several weeks of information disorder had already passed and many people held on to the conspiracy theories and disinformation they initially saw on the online platforms, thus making it hard to tackle the pandemic.

The above discussions show the severe impact that information disorder created in many countries of South and Southeast Asia. It triggered vigilantism and revenge attacks, incited violence against selected communities, fuelled xenophobia, sowed chaos and confusion in social movements, escalated hyper-partisan politics, and challenged democracy. It also shows that the problem gets particularly severe in times of ongoing conflicts and global crisis, like pandemics. In many cases, the impact was immediately visible in the society. But in many cases, the impact was not immediate. Experts view this as a 'slow drip' effect which may gradually inflame tension and change the views of people over time. This can be dangerous for societies with multiple religious, ethnic and cultural groups. It is also viewed that exposure to falsehoods mixed with extremist or partisan views on social media over a long time can skew world views.[187] All these show that the impacts of information disorder have caused sufficient harm and have the potential to escalate in the coming days, if not carefully addressed by all relevant stakeholders.

[187] Thirteenth Parliament of Singapore, op. cit.

Chapter 6

Patterns and Dimensions of Information Disorder in South and Southeast Asia

Discussion on the actors, tactics, and impacts reflects the complexity of information disorder in South and Southeast Asia. The problem is visible in almost all countries of the two regions in varying degrees. Although most of the activities are very country-specific, similarity was seen among the pattern of information disorder in the countries. The influence of information disorder from neighbouring countries and target campaigns from foreign actors was also seen to some extent. This chapter will present an analysis of the overall situation in the two regions and the transnational, regional, and international dimensions of the problem.

6.1 Threat Perception, Legal Frameworks, and Complexities

Governments across the regions have perceived the problem in different ways based on the magnitude of impact in their countries. But one thing is common across most countries of the two regions, i.e., the need to respond to these challenges. Government response involved measures like forming task forces, court rulings, empowering law enforcement agencies, enacting new laws, or applying existing laws. In this regard, Southeast Asia is known to be a region where several news laws have been passed in recent years to tackle disinformation and 'fake news'. Singapore enacted the Protection from Online Falsehoods and Manipulation Act criminalizing fake news and allowing the authorities to remove objectionable online content. As part of Vietnam's cybersecurity law, authorities can demand that social media sites remove false information. Indonesia's Information and Electronic Transactions Act has become its de facto anti-fake news law.[188] Malaysia passed the Anti-Fake News Act in 2018 and Thailand uses the Computer Crime Act to address disinformation.[189] In 2018, the Ministries of Interior, Telecommunications and Information in Cambodia adopted initiatives to punish and penalize those who share 'false information' and block content that creates chaos, damage national defence and security, and incite discrimination or affect national customs and culture.[190] The Philippines has also attempted to take similar initiatives through its Anti-False Content Bill which is pending approval.

[188] "Coronavirus Puts Southeast Asian Anti-Fake News Laws to Test", VOA News, available at https://www.voanews.com/science-health/coronavirus-outbreak/coronavirus-puts-southeast-asian-anti-fake-news-laws-test, accessed on 03 June 2020.

[189] Mong Palatino, "Combatting disinformation in Asia Pacific: Intended—and unintended—consequences", available at https://ifex.org/combatting-disinformation-in-asia-pacific-intended-and-unintended-consequences/, accessed on 18 July 2020.

[190] Khy Sovuthy, "Government to tackle fake news frenzy", *Khmer Times*, available at https://www.khmertimeskh.com/508265/government-to-tackle-fake-news-frenzy/, accessed on 10 August 2020.

In South Asia, the Bangladesh government passed the Digital Security Act, 2018 which addresses issues of information disorder. The act makes it punishable to share content which "creates enmity, hatred or hostility among different classes or communities of the society, or destroys communal harmony, or creates unrest or disorder, or deteriorates or advances to deteriorate the law and order situation."[191] In India, disinformation is addressed through the Indian Penal Code, 1860 under which making, publishing or circulating any statement, rumour or report which may cause fear or alarm to the public, or to any section of the public is a punishable offence.[192] Additionally, in March 2020, India's Ministry of Electronics & Information Technology issued an advisory to all the social media platforms asking to curb false news and misinformation on coronavirus.[193]

While these measures are important for addressing the challenges of information disorder, these have also been subjected to intense debates. In many cases, ambiguities were found in the definitions and interpretations. There have also been instances of misusing it. However, as seen in the analysis of actors, tactics, and impacts, the information disorder in the countries of South and Southeast Asia are very complex and can lead to severe real-life consequences. The scenario is different from many parts of the world. So, in cases where there is an imminent threat of violence and harm against communities, it is important for the government to act. But in this regard, governments face the complex challenge of finding a balance between security and protection of freedom of speech. However, by providing clarity in the laws and preventing misuse, progress can be made in this complex area.

6.2 Influences of the Underlying Social, Political, and Economic Factors

While the analysis of the actors and tactics showed the various forms of information disorder visible in countries of South and Southeast Asia, it is important to understand that in most cases, these are not spontaneous actions. In order to assess the information disorder in the two regions, it is important to take into consideration several social, political and economic factors of the countries along with its media landscape and level of technology adaptation. Here it is seen that most countries of the regions have a colonial history, religious, cultural and ethnic fault lines, and transitioning political environment. The countries are also undergoing a digital transformation that witnessed a high level of internet penetration over a very short span of time. The region has also quickly become home to a large user base of social media companies. Additionally, in many countries, the media has opened up after years of government control. All these factors have given rise to a large number of users and abusers of digital media. The

[191] Legislative and Parliamentary Affairs Division, *Digital Security Act, 2018*, Dhaka: Ministry of Law, Justice and Parliamentary Affairs, Government of the People's Republic of Bangladesh, 2019.
[192] Khushbu Jain and Brijesh Singh, "Disinformation in times of a pandemic, and the laws around it", *The Economic Times*, 03 April 2020.
[193] Ibid.

availability of low paid workforce in most countries also made these regions a hub for troll industries and disinformation syndicates. Thus, the presence of several actors with varied interests to influence the information ecosystem and the underlying political, social, religious and ethnic fault lines in the countries, made it a fertile ground for information disorder.

Myanmar can be used as a case study in this regard. The combination of the mentioned factors has played a role in giving rise to hate and intolerance in the information ecosystem and ultimately leading to oppression of its minorities, particularly the Rohingyas. After emerging from decades of long military rule, Myanmar went through several transitions. A country which was once one of the least connected regions of the world with only 1.1 per cent population using the internet according to the International Telecommunication Union (ITU), witnessed rapid penetration of the internet after the deregulation of telecommunications by the then government in 2013.[194] The price of SIM cards dropped from more than US$200 to as little as US$2 and by 2016, nearly half the population had mobile phone subscriptions and smartphones with internet access.[195] In this context, Facebook became so popular that it became synonymous with the internet. Mobile phone operators also banked on this opportunity by offering deals which included using Facebook without data charge. The free basic initiative to make the internet available to the people in developing countries by Facebook also contributed to its large user base. All these factors made digital tools very much available, accessible, and impactful for the hardliner religious actors to spread hate speech and disinformation. The low digital literacy level and pre-existing intolerance towards ethnic and religious minorities also played a key role here. All these gave rise to an environment where disinformation and hate speech reigned unchecked for a time long enough to cause unprecedented damage to the Rohingya community.

The influence of these underlying conditions was also prominent in India. Researchers opine that the emergence of problems like rumours-based vigilantism and escalation of hate speech against minority communities need to be studied from the broader context of a society experiencing technological transformation. In this regard, this paper explores the underlying conditions of India based on the study of Shakuntala Banaji and Ram Bhat of the London School of Economics and Political Science.[196] Their study shows that while smartphone usage has been on the rise in India since 2013, the country witnessed a radical shift in internet usage in 2016. In this period, free and unlimited data were provided to all subscribers by a leading telecommunication company and the data tariff of other companies were driven down significantly. This allowed the internet to spread from the urban centres to peripheral regions very fast. Parallelly, smartphones also became very affordable due to the import of low-cost Chinese smartphones and also the production of relatively low-cost phones by a few Indian manufacturers. It is

[194] Steve Stecklow, op. cit.
[195] Ibid.
[196] Shakuntala Banaji and Ram Bhat, op. cit.

estimated that in a span of ten years, the price of smartphones came down to 16 per cent. Due to the proliferation of smartphones and affordable data packages, communication services became widely popular and India became one of the most significant markets for online platforms like YouTube, Facebook, Facebook Messenger, TikTok, WhatsApp, Instagram, Telegram, ShareChat, SHAREit, Zapya, etc. As India is a country with relatively low levels of textual literacy, the ability to exchange audio-visual content was seen as a liberating experience for many users. While this development helped in getting a large segment of the population connected to the internet, it also creates scopes for abuse. This transitioning environment along with prejudiced ideological positions and discriminatory beliefs rooted in social, historical, and political fault lines, made India a fertile ground for information disorder.

Bangladesh also witnessed a similar influence of historical, geopolitical, and social factors. Md. Sayeed Al-Zaman described this phenomenon using William F. Ogburn's concept of 'cultural lag'. His study indicates that communication technology as a material culture has been penetrating and revolutionizing Bangladesh society, but the pre-existing belief and morality cannot cope up with the material change, thus resisting the process. Digital disinformation can be seen as an expression of this propensity.[197] Moreover, like many other countries of the region, the rapid penetration of the internet to a large user base with low information literacy is also a major factor in Bangladesh. Similar to Myanmar and India, the country also witnessed wide popularity of free services offered by social media companies as well as telecommunication operators. However, after many years of use and abuse, in a recent decision in July 2020, Bangladesh Telecommunication Regulatory Commission (BTRC) ordered mobile phone operators to ban free internet for social media in order to ensure safe internet. One of the reasons stated was that "'some dishonest people' were carrying out 'unnecessary' criminal activities on social media by using the free services."[198] While this might bring temporary discomfort to a large user base, it is also important to note how such services were used in amplifying information disorder.

The similar influence of underlying social, political, religious, and historical factors was visible in other countries too. Sri Lanka and Thailand need special mention in this regard as it was reflected in the incidents of hate speech targeting minority communities. These factors along with economic factors like availability of low paid workforce were visible in the Philippines and Indonesia. The combination of all these factors indicates why South and Southeast Asia witnessed such a large-scale information disorder campaign in the past few years.

[197] Md. Sayeed Al-Zaman, "Digital Disinformation and Communalism in Bangladesh" *China Media Research*, Volume 15, Issue 2, 2019.

[198] "Bangladesh bans free internet for social media to stop 'unhealthy' competition", *bdnews24.com*, 18 July 2020.

6.3 Transnational Implications of Information Disorder

Through the analysis of the actors, tactics, and impacts in South and Southeast Asia, it is seen that there is some kind of transnational resemblance and connection among selected countries of the two regions. In many cases, it is seen that a particular incident in a country triggers the rapid sharing of misinformation, disinformation, and hate speech in the neighbouring countries with the same religious and ethnic groups. This can be seen in three forms.

Firstly, the similarity in content and narratives among actors of the same religion. In this regard, a common pattern was seen among the online activities of hard-line Buddhist groups and their leaders in Sri Lanka, Myanmar, and Thailand. Actors of the three countries abused the power of social media to propagate hate speech and fear monger Buddhists against the minority communities, particularly the Muslims. In-depth interviews with fact-checkers and analysis of the contents showed that there is a common trend in the anti-Muslim hate speech and images shared by the Buddhist hardliners in Myanmar and Sri Lanka, for example, Muslim aggression in their communities, Muslim conspiracy to take over Buddhist dominance through rapid birth rate, call for boycotting Muslim business, dehumanization of Muslims, etc. Previously, studies by scholars have found stark similarities in the strategies of Ma Ba Tha Movement in Myanmar and Bodu Bala Sena in Sri Lanka, like their extensive use of social media.[199] Although the movements spearheaded by Ma Ba Tha and BBS have historical precursors, the developments in communication technology made social media a new, effective vehicle for outreach and mobilization. While anti-Muslim sentiments existed before, the spread of ICT made it much more compelling and convincing with images and videos and it could be instantly spread to the whole country.[200] Alongside these two countries, Thailand also witnessed a similar kind of anti-Muslim hate speech and fear-mongering over social media. The case of monk Aphichat Punnajanto needs to be considered in this regard. Although each country has its history, causes, and instigators, there seems to be a common narrative among the actors of all three countries, i.e., Buddhism is somehow under threat and that Islam and Muslims are trying to take over their country.[201]

Secondly, the similarity in narrative among actors of different religious communities against a common target group is visible. In this regard, it was seen that Islamophobic narrative from Buddhist hardliner groups in Sri Lanka and Myanmar was

[199] Usaid Siddiqui, *Muslim Minorities in Peril: The Rise of Buddhist Violence in Asia*, Doha: Al Jazeera Centre for Studies, 2016 and Vishal Arora, "Connecting the Dots on Buddhist Fundamentalism", *The Diplomat*, 30 May 2014.

[200] Camilla Orjuela, "Countering Buddhist Radicalisation: Emerging Peace Movements in Myanmar and Sri Lanka", *Third World Quarterly*, Volume 41, Issue 1, 2020. pp. 133-150.

[201] AFP, "Rise of violent Buddhist rhetoric in Asia defies stereotypes", available at https://www.bangkokpost.com/world/1426722/rise-of-violent-buddhist-rhetoric-in-asia-defies-stereotypes, accessed on 01 July 2020 and Deutsche Welle, "Buddhists fan flames of Islamophobia in Southeast Asia", available at https://www.dw.com/en/buddhists-fan-flames-of-islamophobia-in-southeast-asia/a-43158407, accessed on 01 July 2020.

used in the disinformation and hate speech against the Muslims in India. For example, the study of Equality Labs found the repeated celebration of the Buddhist monk named Ashin Wirathu among the Hindu radicals on Twitter and Facebook. Messages of his oppression on the Rohingya community and boycott of the Muslims and their business in Myanmar were widely circulated.[202] The actors are reported to have intentionally used the misspelled hashtag #Virathu to avoid detection by the algorithms. The previous study of Equality Labs on hate speech in India found that 6 per cent of Islamophobic posts studied were specifically anti-Rohingya. The posts show the use of misleading and fabricated images used to frame the Rohingyas and fear monger Hindus against them.[203] According to the report, "usually these stories claim that Rohingya refugees are swarming India; that they are swaying elections, obtaining false identification, or that they have married 'local girls' who are giving birth to multiple children very quickly and swelling the population of Muslims in India."[204] Additionally, disinformation from Sri Lanka was also used in India. For example, the false reports of the Muslim doctor sterilizing Buddhist women in Sri Lanka were also shared by social media users in India where some posts called for inquiring into whether Hindus were being sterilized in India as well.[205]

Thirdly, a transnational impact is seen in support for the victims or target groups by similar religious communities in nearby countries. For example, during the peak of Rohingya oppression in 2012 and 2016-2017, misinformation was widely circulated in Bangladesh using false context, and fabricated photos in support of the Rohingyas and disinformation was spread against the Buddhists. While it is understandable that Bangladesh being a Muslim majority country will resonate with the sufferings of the Rohingya community facing atrocity, it was seen that the widely circulated images in social media were not always related to this particular incident. The transnational effect surrounding the Rohingya issue was also seen in Singapore. Inflammatory comments having Islamophobic overtones targeting the Rohingya community from seemingly Myanmar-based user accounts about articles written on the Rohingya issue in Singapore created an online backlash from Singaporean Muslims, resulting in heightened tensions along religious and ethnic lines between the users.[206] The transnational implication was also seen in Indonesia where the Muslim Cyber Army disseminated misinformation and malinformation surrounding the persecution of Muslims in Myanmar in a manner that is domestically relatable.[207]

Similarly, it is seen that incidents targeting Muslims in India or related issues, also have a ripple effect among many Muslim groups in Sri Lanka, Bangladesh,

[202] T. Soundararajan et al., 2020, op. cit.
[203] T. Soundararajan et al., 2019, op. cit.
[204] Ibid.
[205] AFP Sri Lanka, op. cit.
[206] Thirteenth Parliament of Singapore, op. cit.
[207] Kate Lamb, op. cit.

Pakistan, and the Maldives. Disinformation campaigns are often spread surrounding it.[208] Alternatively, it was also seen that incidents surrounding the Hindu community in Bangladesh are shared in messaging services and local online news portals in India. In this process, sometimes a regular incident is also reportedly shared as disinformation with a communal spin.[209] It is important to be vigilant of such activities as it can harm goodwill and create tension among the people of bordering countries.[210]

These forms of transnational effect of information disorder indicate a growing problem which can have several implications in the bilateral relations between countries and harmony among the religious communities within the country. Thus, close observation of online activities in the neighbouring countries, particularly related to religious communities is very important to assess the information disorder in one's own country.

6.4 Resemblance of Information Disorder in South and Southeast Asia

The countries of South and Southeast Asia are closely connected in the real and virtual world within their region and also between the two regions. Most countries share similar ethnic and religious fault lines which have influenced information disorder nationally and also transnationally, as seen in the above discussion. Increasingly, there are concerns regarding the regional implications of the problem. It is speculated that the financially motivated actors available in many countries of the region can easily be used to engage in disinformation activities against a target country. Such acts can be very deceiving owing to similarities in the culture and language in the region. In this regard, the Singapore Parliament's Select Committee on Deliberate Online Falsehoods assessed that the developing disinformation capabilities in the region like for-profit syndicates, bot armies, and data-driven political consultants can be repurposed and deployed against Singapore.[211] Thus, it is important for the countries to be vigilant of such activities in the region.

Moreover, in the hyper-connected world of the internet, ideas and skills are transferred very swiftly. It is seen that the tactic used in different campaigns in the region are quickly adopted by actors in other countries. Inspiration is often taken from the contents. Those are then contextualized and remade for the target audience of respective countries. For example, the resemblance in the anti-Muslim narrative and images shared in Myanmar, Thailand, Sri Lanka, and India. Such practice was also widely seen in the early days of the COVID-19 pandemic. For example, rumours of a new-born child with

[208] Based on an interview with a researcher in Delhi, India on 09 May 2020.

[209] See more at Qadaruddin Shishir, "Government Officials Land Rescue Operation Reported with Communal Spin", Boom, available at https://www.boombd.com/fact-file/govt-officials-land-rescue-operation-reported-with-communal-spin-8873?fbclid=IwAR2DGauoFyuS9O7v3G5UZ_QMULHflyu60Y08TmzrpBYGU9XyeB162ZS1tc4, accessed on 22 July 2020.

[210] Based on an interview with Mahbub Roni, Co-Founder and Secretary, BD FactCheck on 22 July 2020.

[211] Thirteenth Parliament of Singapore, op. cit.

Anencephaly prescribing coronavirus remedy right after birth circulated among social media users in Bangladesh and India, however, the story was contextualized along separate religious lines.[212] The similarity is also seen in the impacts. For example, the child kidnapping rumours in India, Bangladesh in Indonesia leading to mob lynching. Thus, it is important to keep track of disinformation campaigns in the two regions and assess the possibilities of replication. Analysis of the disinformation scenarios in countries of the region with similar socio-cultural contexts can be effective in predicting or assessing potential risks and challenges in one's own country.

6.5 Alarming Rise of Radical Narratives and Security Implications for the Regions

The digital platforms have been largely held responsible for creating scopes of online radicalization. Social media has been extensively used by ISIS and its affiliated organizations.[213] The Christchurch incident in New Zealand added an entirely new dimension to promoting terrorist acts through these platforms. However, due to the collaborated efforts of big tech companies, governments, law enforcement agencies, and security experts, it has been possible to contain a portion of the publicly available extremist content coming from Islamist militant groups. However, many of the groups are still active in online platforms. Additionally, in the region of South and Southeast Asia, the growing trend of hard-line Hindu and Buddhist groups using social media to incite violence, hatred, and radical narratives need to be closely studied. It is seen that besides explicitly promoting extremist narratives, the groups have also resorted to a subtle and reframed version of the radical content. Such contents fall less into the extremist criteria and lean more towards disinformation and hate speech thus, often bypass the filters of the social media platforms, fact-checkers and monitoring agencies. It is important to revaluate the policy of the stakeholders in this regard. This problem also has security implications for the regions as the intensifying anti-Muslim hate messages and disinformation can trigger reactive online campaigns from hard-line Muslim groups in the countries. It can also influence campaigns in countries where Muslims are the majority and Hindus and Buddhists are minorities.

6.6 Information Disorder from Foreign Sources: Hostile Information Campaign

While most of the origins of information disorder in South and Southeast Asia are domestic, disinformation from foreign states was recently recorded in few countries, namely, India, Singapore, and Thailand.

[212] Alt News, "Photos, video of baby who died due to birth defect viral as newborn talking after birth", available at https://www.altnews.in/images-video-abnormal-newborn-child-started-speaking-after-birth-false-claim/, accessed on 03 June 2020.

[213] Brenden I. Koerner, "Why ISIS Is Winning the Social Media War", Wired, available at https://www.wired.com/2016/03/isis-winning-social-media-war-heres-beat/, accessed on 12 October 2020.

In April 2019, Facebook reported that it removed a network of 103 pages, groups, and accounts originating from Pakistan for engaging in coordinated inauthentic behaviour on Facebook and Instagram. Facebook's investigation of the network found that it was linked to employees of the Inter-Services Public Relations (ISPR) of the Pakistani military.[214] The report revealed that the individuals behind this activity used fake accounts to operate different fan pages, paid for Facebook advertisements to boost their reach and frequently posted about local and political news including topics like the Indian government, political leaders, and military.[215] Analysis of the Atlantic Council's Digital Forensic Research Lab showed that during the heightened tensions between Pakistan and India surrounding the Pulwama attack, many of these accounts were reported to have inflamed tensions with India.[216] In such circumstances, the use of computational propaganda to advance the political and strategic objectives of a country can have an adverse psychological impact.[217] This adds a new dimension to the study of conflicts between the two South Asian neighbours.

Another case of Hostile Information Campaigns was reported in Thailand by RSIS's policy report on Foreign Interference in Asia.[218] The report referred to Facebook's disclosure on the incident. As part of Facebook's investigation into suspected Thailand-linked accounts, in July 2019 the company disclosed that it had removed 12 Facebook accounts and 10 Facebook pages for engaging in coordinated inauthentic behaviour that originated in Thailand and focused primarily on Thailand and the US.[219] According to the report, the network used fake accounts to create fictitious personas and run pages, increase engagement, disseminate content, and also drove people to off-platform blogs posing as news outlets. These accounts and pages frequently shared divisive narratives and comments on topics including Thai politics, geopolitical issues like US-China relations, protests in Hong Kong, and criticism of democracy activists in Thailand. Facebook found that some of this activity was linked to an individual based in Thailand associated with a Russian government-funded journal based in Moscow. The investigation into some of these pages by Digital Forensic Research Lab revealed that the pages boosted hostile narratives and positioned themselves as 'alternative media' that countered Western stances on international issues, with a particular emphasis on Thailand.[220]

---

[214] Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior and Spam From India and Pakistan", Facebook Newsroom, available at https://about.fb.com/news/2019/04/cib-and-spam-from-india-pakistan/, accessed on 15 June 2020.

[215] Ibid.

[216] Digital Forensic Research Lab, "Pakistan Army's Covert Social Network", available at https://medium.com/dfrlab/pakistan-armys-covert-social-network-23ce90feb0d0, accessed on 15 May 2020.

[217] See more at Shaurya Karanbir Gurung, "Defence ministry approves information warfare branch for Indian army", *The Economic Times*, 09 March 2019.

[218] Muhammad Faizal Bin Abdul Rahman et al., *Cases of Foreign Interference in Asia*, op. cit.

[219] Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior in Thailand, Russia, Ukraine and Honduras", available at https://about.fb.com/news/2019/07/removing-cib-thailand-russia-ukraine-honduras/, accessed on 19 June 2020.

[220] Digital Forensic Research Lab, "Facebook Takes Down Inauthentic Pages with Connections to Thailand", available at https://medium.com/dfrlab/facebook-takes-down-inauthentic-pages-with-connections-to-thailand-7dbf331f5ba5, accessed on 19 June 2020.

Singapore Parliament's Select Committee on Deliberate Online Falsehoods reported that Singapore has been the subject of foreign state-sponsored disinformation operations which included "a State using news articles and social media to influence the minds of segments of the Singapore population, and to legitimize the State's actions in the international sphere."[221] Highlighting the risk of such actions, the committee chairman Charles Chong stated that online falsehoods are "pervasive and can affect different aspects of our country: national security, racial harmony, democratic processes, social cohesion and trust in public institutions".[222] While hostile information campaigns by foreign actors were conducted in small scales so far, it is perceived that this can magnify into big threats. Edwin Tong, Senior Minister of State for Law of Singapore, stated that "In this battlefield, Singapore, an open, democratic, digitally-connected and diverse country, is especially vulnerable. We are a young country with sensitive fault-lines that foreign actors can exploit to foment distrust and ill-will among our various communities."[223]

Hostile information operations in cyberspace by foreign states is considered to be a national security concern as it can undermine a state's sovereignty by interfering in domestic activities through deceptive online tactics. Such actions can sow discord among communities, weaken the resilience of people, undermine policies, discredit public institutions, and even influence elections. This is a growing concern among states in the West. As few countries in South and Southeast Asia have already experienced some forms of information operations, it is important for the governments to closely monitor such possibilities and take precautionary measures.

6.7 Influence of International Information Campaigns

The information ecosystem of South and Southeast Asia has also been affected by the information campaigns on the global stage. The 'Infodemic' and war of words among top government officials surrounding the COVID-19 pandemic largely contributed to the information disorder in the present time. The US State Department's Global Engagement Center (GEC) has alleged that Russian, Chinese, and Iranian state information operations are converging around the same disinformation narrative themes about COVID-19. According to the US special envoy and coordinator of the GEC, Lea Gabrielle's Briefing on Disinformation and Propaganda Related to COVID-19, China's malign disinformation falsely blamed the US as the origin of the virus and made effort to turn the crisis into a news story

[221] Thirteenth Parliament of Singapore, op. cit.

[222] Yasmine Yahya, "Select Committee on fake news: Singapore a target of hostile info campaigns", *The Straits Times*, available at https://www.straitstimes.com/politics/spore-a-target-of-hostile-info-campaigns, accessed on 20 June 2020.

[223] Adrian Lim, "Parliament: 'Curious' spike in online comments critical of S'pore during dispute with Malaysia, says Edwin Tong", *The Strait Times*, available at https://www.straitstimes.com/politics/parliament-curious-spike-in-online-comments-critical-of-spore-during-dispute-with-malaysia, accessed on 20 June 2020.

highlighting supremacy of the Chinese Communist Party in handling the health crisis.[224]

On the other spectrum, China's Ministry of Foreign Affairs published an article rebutting 24 claims from the US, including those calling the novel coronavirus 'the Chinese virus' or 'Wuhan virus' and that the virus is artificially made.[225] Additionally, China is reported to have engaged in a 'Wolf Warrior Diplomacy' primarily through social media platforms to push forward its narratives. All these promoted chaos and uncertainty surrounding the pandemic. These also undermined the efforts of health organizations that are in charge of disseminating accurate information about the virus. It also influenced the thoughts of general people and made them suspicious. Here it can be seen that although countries of South and Southeast Asia were not directly engaged in this activity, the people of these regions were still influenced by such campaigns.

The analysis of the pattern of information disorder in South and Southeast Asia helps to understand why the issue developed and how it escalated in the two regions. It reveals the role of underlying factors which make the information disorder different from other parts. It highlights the threat perception by the government and the challenges related to it. It also brings forward the transnational, regional, and international dimensions to the problem and the security implications. All these show that the nature of the problem is highly complex and difficult for countries to tackle alone. Thus, collaborated measures need to be thought of. The following chapter attempts to put forward some feasible options for the countries to consider.

[224] U.S Department of State, "Briefing on Disinformation and Propaganda Related to COVID-19", available at https://www.state.gov/briefing-with-special-envoy-lea-gabrielle-global-engagement-center-on-disinformation-and-propaganda-related-to-covid-19/, accessed on 02 June 2020.

[225] Ministry of Foreign Affairs, China, "The China-related Lies and Facts about the New Coronary Pneumonia Epidemic", available at https://www.fmprc.gov.cn/web/ziliao_674904/zt_674979/dnzt_674981/qtzt/kjgzbdfyyq_699171/t1777471.shtml, accessed on 02 June 2020.

Chapter 7

Way Forward

The problems surrounding information disorder has affected several aspects of the society and state. As discussed in the preceding chapters, information disorder has implications for both national and human security. It can inflict severe damages in real life including harming individuals, oppressing communities, destabilizing religious harmony and social cohesion, impacting public institutions, and law and order situation of the country. Almost all countries of South and Southeast Asia are experiencing the problem and taking different measures to address this issue. While individual countries have taken different measures to tackle the problem within their jurisdiction, it is seen that many a time the problem is transnational and global in nature, which requires coordinated approaches among the countries and the related stakeholders. This chapter puts forward some of the policy recommendations that can be taken individually by the countries and also through regional and global collaboration.

7.1 Promote Research and Documentation of the Problem

Documentation of the problem and research on it is key to addressing this challenge of information disorder. But during this study, it was seen that there are limited scholarly works on the topic and inadequate documentation of incidents in many countries of both the regions. While it is understandable that the countries experience disproportionate levels of the problem, there were limited resources even in countries where the impact is significant. Thus, more academic efforts in this area encompassing both the offline and online impacts would help to address the problem.

It is also important to contextualize definitions and frameworks based on the experience of the countries. The social, historical, and cultural context of the actors and their target audience is vital in understanding the different types of information disorder and assess possible impacts. Research and analysis from local experts with an understanding of the language and psychology of different communities are crucial in this regard. Besides studying individual countries, it is also important to research on the kinds of information disorder that impacts similar communities in the neighbouring countries to understand the transnational pattern of the problem.

It is observed that leading universities and think tanks in many parts of the world are increasingly undertaking projects and researches on related issues. Separate internet institutes and digital labs are also being formed to conduct in-

depth studies on this topic. South and Southeast Asia have very limited places where such a study is conducted. It is important for countries to recognize this fast-evolving field and take appropriate initiatives.

7.2 Establish and Promote Fact-Checking Initiatives

The role of fact-checking organizations is vital in addressing information disorder. By analyzing the disputed or suspected content, the fact-checkers provide an objective analysis of the information. Leading social media platforms work with these organizations to verify content in their platforms. Based on the report from fact-checking organizations, the content is either removed, a warning label is attached to aware the users or, it is made less trending on the newsfeed. However, the number of fact-checking organizations are limited in South Asia and Southeast Asia. Among the limited number, only a few are part of the International Fact-Checking Network (IFCN), which includes, Rappler and Vera Files in the Philippines, Boom in India and Tirto.id in Indonesia.[226] IFCN is a global forum of fact-checkers by the Poynter Institute. Companies like Facebook only work with these verified organizations. However, many countries of South and Southeast Asia do not have an IFCN-certified organization.

While it is understandable that social media companies need to ensure that the organizations moderating content on their platforms are trusted and maintains a high professional standard, various constraints limit fact-checking organizations in this part of the world from meeting the requirements. For example, difficulty in registration, lack of digital archives, financial resources, technological innovation, skilled manpower, etc. Also, many organizations work on a part-time basis with few volunteers or employees, so it is difficult for them to respond to the fast-paced circulation of disputed content. But these local organizations are fundamental to understanding and analyzing local content. Thus, it is important for countries to facilitate fact-checking initiatives and promote their works so that the debunked content is widely viewed. It is also important for social media companies to find ways to engage more with the local fact-checkers.

Bangladesh has only two established fact-checking initiatives i.e., BD FactCheck and Fact Watch. While these two organizations are actively working within their limited resources and have gained the trust of online users, the scale of their operation is inadequate to cater to the needs of such a large user base in the country. Analyzing several hundred contents of information disorder circulating

[226] Wataru Suzuki, "Facebook's Fact-Checking in Asia Faces Challenges", available at https://asia.nikkei.com/Business/Business-trends/Facebooks-factchecking-in-Asia-faces-challenges, accessed on 20 June 2020.

every day is also difficult for the two organizations. Thus, it is important to facilitate the expansion of these operations.

Also, in most cases, the content reviewed by the organizations are usually the ones circulating in the leading social media sites and big online news portals. But it is also important to bring the online news portals catering to sub-urban regions under its purview. As discussed in Chapter 3, online news portals focusing on different localities or specific regions within the country can be one of the major actors of disinformation and rumours. Thus, it is important to include journalists and academics from different parts of the country to address the disinformation and rumours faced by the rapidly growing user base outside the big cities.

Besides scaling up the two initiatives, new initiatives with a sustainable business model also need to be considered. In this regard, the trusted mainstream news and media agencies can be focused on. In many countries, leading media houses have their own fact-checking initiatives which they use for their in-house analysis and also disseminate the reports publicly. As these organizations already have skilled and experienced professionals, fact-checking by them would be more efficient and credible. Also, it would make the organizations more responsible for the contents that are shared by them, both online and through traditional outlets. The inconsistency of the editorial policies in the online versions of traditional media was identified in Chapter 3. Thus, the combination of expanding existing initiatives, launching new initiatives and insisting social media companies to work with local fact-checkers, can be an effective way forward for the countries.

7.3 Promote Strong and Trusted Government Communication and Collaboration Among Agencies

Strong and trusted communication from the government is crucial in addressing the information disorder. As discussed in Chapter 5, many a time disinformation is spread by sourcing government officials and agencies which creates confusion among the citizens. Misinformation is also spread surrounding a foreign country and its leaders that can risk harming bilateral relations. In regard to such incidents, it is important for the governments to have strong communication platforms which people can easily access and get clarifications. In this regard, the practices of Singapore and Malaysia can be taken into consideration.

In Singapore, the Public Communications Division of the Ministry of Communications and Information manages a web portal called 'Factually' which aims to "clarify widespread or common misperceptions of government policy or inaccurate assertions on matters of public concern that can harm Singapore's

social fabric."[227] The website regularly provides 'corrections and clarifications' on different types of information disorder visible across media platforms in the country. In Malaysia, the Communications and Multimedia Ministry (KKMM) in collaboration with the Malaysian Communications and Multimedia Commission (MCMC) operates 'sebenarnya.my' to prevent the spread of 'false news' which may affect national harmony and security.[228] This is an information verification portal where the public can verify suspicious content that they receive through social media, short messaging services, blogs, or websites. It also serves as an avenue for ministries and governments to debunk disinformation or make clarifications on issues that have become viral. The portal also works as an archive for fake news involving the public's interest and the nation.[229] Both these initiatives have been effective in addressing the challenge of information disorder. The government of Bangladesh as well as other countries in the regions can also explore the possibilities of taking such initiatives in collaboration with their affiliated ministries.

In this regard, collaboration among agencies is crucial. It is seen that information, ICT and telecommunication are often separately looked after by different government bodies, for example, the Ministry of Information, Ministry of Posts, Telecommunications and Information Technology and ICT Division in Bangladesh. But information disorder is a complex phenomenon which is associated with all three and also the Ministry of Home Affairs. In this regard, effective collaboration and capacity building is required between the relevant stakeholders of the government to address this challenge.

7.4 Proactive and Issue Focused Measures

It is important for governments and civil society organizations to take proactive measures focused on specific issues. Safeguarding elections from information disorder is a big issue worldwide. As elections campaigns have shifted to the digital platforms, it is important to ensure that the actors abide by the regulations of the election commission in cyberspace. The abuse of digital platforms by cyber troops, trolls, bots, and fake news syndicates have the potential to manipulate opinion. Such practices were visible in varying scales in many countries of South and Southeast Asia. So, it is important for countries to strengthen

[227] Ministry of Communications and Information, Singapore, available at https://www.mci.gov.sg/pressroom/news-and-stories/pressroom/2018/2/mcis-response-to-pq-on-gov-factually, access on 29 June 2020.

[228] Malaysian Communications and Multimedia Commission (MCMC), *Sebenarnya.my Portal Launched, In a Battle Against False News*, Cyberjaya: Malaysian Communications and Multimedia Commission (MCMC), 2017.

[229] Paige Occenola, "Fake News and Freedom of Expression in Southeast Asia", available at https://www.rappler.com/technology/social-media/188573-fake-news-freedom-expression-southeast-asia, accessed on 29 June 2020.

their precautionary measures during such a crucial period. It is also necessary to vigilant of hostile information campaigns from foreign actors. While only selected countries have reported such activities, it is important to closely monitor global and regional activities in this regard.

It is also important to take pre-emptive measures when there is tension, conflict, or violence along communal lines in the region so that it does not have a ripple effect among similar communities in other countries. Additionally, it is essential to assess misinformation, disinformation, and rumour related to issues sensitive to the regions, for example, child kidnapping, sterilization, vaccination, refugees, religious defamation, etc. The transnational and regional implications of such campaigns need to be considered.

7.5 Updating Legal Frameworks and Approving Draft Policies

The digital world is subject to constant changes. New and innovative tactics of information disorder are adopted very frequently. It is important for the governments to update their legal frameworks to address these technologically sophisticated tactics. In doing so, it is also important to address the ambiguous sections of the existing laws to prevent misuse. Additionally, it is necessary to consider the inclusion of provisions that address foreign interference conducted through online mediums as this is a growing national security concern in many countries.

It is also essential to update and approve the pending policies. For example, in Bangladesh, the National Online Mass-Media Policy was drafted in 2015. In 2017, the cabinet approved the draft and incorporated a proposal for setting up a National Broadcast Commission which will facilitate the operation of online mass media.[230] But there have been significant changes in the online information ecosystem over the past five years, so it is important to update the draft and take it to the next stages of approval.

7.6 Promoting Resilience and Critical Media Literacy

It is seen that the majority cases of misinformation which contribute to the formation of mobs, lynching, and other discriminatory violence, are largely spread based on prejudice, ideological positions, and discriminatory beliefs. So, it is important to look for an effective mechanism which addresses the underlying factors. Promoting resilience and inter-faith dialogue is crucial in this regard. In

[230] "Registration must for online mass media", *The Daily Star*, 20 June 2017.

order to effectively and sustainably address information disorder, it is also important to sensitize people about the problem. They need to be made conscious to critically think about why they are seeing what they are seeing on their newsfeed and what might be the motivation behind certain contents. They also need to be made aware of responsible information sharing and the consequences that misinformation can lead to, in order to break the chain of circulation and reduce impact. In this regard, critical media literacy is extremely crucial for developing knowledge and promoting responsible consumption of information. While there has been some intervention in this regard in a few countries, it is seen that such programs are often limited within big cities and tertiary level students. Initiatives need to be taken to roll out such programs both through formal education and informal awareness-building activities to reach the mass level, particularly the semi-urban and rural regions within the countries. Tailored messaging and communication strategies for various demographics is crucial in crafting effective media literacy and promoting tolerance.

7.7 Regional Collaboration

The transnational and regional implications of information disorder were reflected in Chapter 6 of this paper. It is difficult for individual countries to address such issues in silo. In this regard, the countries of South Asia and Southeast Asia can work within their regions and also collaborate between the two regions in areas of mutual agreement and by respecting the cultural and political landscape. A regional mechanism developed with the collaboration of government officials, fact-checkers, security analysts, academics, journalists, and think-tanks can be an effective way forward. Regional approaches are already in practice in some parts of the world, like the EU East StratCom Taskforce formed in 2015.[231] The Association of Southeast Asian Nations (ASEAN) has also taken initiatives in this direction. In the ASEAN Ministers Responsible for Information (AMRI) Roundtable Discussion on Fake News and Communicating the Right Information, 2017, it was observed that the proliferation of fake news could cause hate and conflict that can potentially undermine social cohesiveness.[232] In the 14th Conference of the ASEAN Ministers Responsible for Information (14th AMRI), 2018, a 'Framework and Joint Declaration to Minimise the Harmful Effects of Fake News' was adopted. The framework highlighted education and awareness, detection and response and,

[231] The Task Force was set up to address Russia's disinformation campaigns. More information is available at European External Action Service, https://eeas.europa.eu/headquarters/headquarters-homepage/2116/-questions-and-answers-about-the-east-stratcom-task-force_en.

[232] ASEAN, "ASEAN to Cooperate on Fighting Fake News in the Region", available at https://asean.org/asean-to-cooperate-on-fighting-fake-news-in-the-region/, accessed on 28 June 2018.

norms and guidelines as strategies for tackling the issues surrounding fake news.[233] The joint declaration reaffirmed the need to address the proliferation of fake news and its negative impact to ensure that the internet remains a reliable source of information and a safe space for all users.[234] The declaration also highlighted the need to continue the sharing of best practices among the ASEAN Member States on sensitizing citizens to the harmful effects of fake news.

Such regional approaches can also be considered in South Asia. In this regard, the Bay of Bengal Initiative for Multi-Sectoral Technical and Economic Cooperation (BIMSTEC) can be considered. The organization came into being in 1997 but in recent times, it has created momentum by initiating several activities for the member countries.[235] BIMSTEC has initiated Track 1.5 BIMSTEC Security Dialogue Forum and Think Tanks Dialogue on Regional Security among the member states. Technology is already a priority area of the organization and the recent incorporation of the security paves the way for enhanced cooperation in cyberspace. Thus, BIMSTEC can take initiatives to engage with the member states to address the menace of information disorder. This can be done in collaboration with the relevant stakeholders within the countries. It would facilitate the sharing of best practices, identifying patterns, and conducting collaborated researches on the topic. Cooperation within BIMSTEC will also allow for intra-regional cooperation between South and Southeast Asia. Additionally, coordination and collaboration among the regional blocs will assist in raising collective voices in the global negotiations for content moderation and ensuring a safer internet.

7.8 Database on Regional Issues

Creating a database of contents of information disorder prevalent in South and Southeast Asia can also be initiated by the think tanks of the countries. This can help researchers, academics, fact-checkers, and policymakers understand the current trend of the problem and take necessary measures in respective countries. One of the biggest challenges the researchers and fact-checkers face in analyzing content is language barriers. A collaborated database with translations can be a useful resource for studying this topic. In this regard, the COVID-19 database formed by the International Fact-Checking Network (IFCN) can be used as a reference. The IFCN database is a collection of debunked misinformation published across 70 countries and in at least 40 languages.[236] Similar initiatives regarding

[233] 14th Conference of the ASEAN Ministers Responsible for Information (14th AMRI), *Framework and Joint Declaration to Minimise the Harmful Effects of Fake News*, Singapore: ASEAN, 2018.
[234] Ibid.
[235] It constitutes seven member countries: Bangladesh, Bhutan, India, Nepal, Sri Lanka from South Asia; and Myanmar and Thailand from Southeast Asia.
[236] Poynter, "The CoronaVirusFacts/DatosCoronaVirus Alliance Database", available at https://www.poynter.

hate messages, rumours, misinformation, and disinformation need to be created for the countries of South and Southeast Asia.

7.9 Addressing Language Barriers

The big techs depend on both manual and automated content moderation. While human moderators play an important role in analyzing and contextualizing content, it is difficult for humans to filter through the billions of contents uploaded every day. Automated content detection is instrumental in this regard. A good natural language processing system can serve as a faster and efficient filter. But this is where the barrier of language comes in. Most of these technologies work in English. Algorithms already built will be able to work reasonably well within this language family. But the natural language processing does not work the same way for most languages in the Global South, particularly the Indo-Aryan languages in South Asia whose alphabets, vocabularies, syntaxes, and morphologies are very different from other branches and there are insufficient statistical resources required for prompt analysis.[237]

Also, when it comes to translating disputed texts like hate speech to a different language for analysis, a lot is lost in the translation and the resulting version does not necessarily capture the sentiment of the original post. Thus, many contents pass through these cracks of the detection system. In order to effectively solve the problem of information disorder, addressing the language barrier is crucial. The big techs need to be urged to work in this field.

7.10 Participation in Global Frameworks for Content Governance

While digital media platforms have depended on self-regulation, several instances of abuse of these platforms across the world, particularly in South and Southeast Asia indicate that these are inadequate and companies need to do more to keep their platforms free of disinformation, hate speech, and violence.

This has led to the rise of discussions regarding global governance of content moderation. The UN's Roadmap for Digital Cooperation acknowledged the need for content governance of harmful online content. This report of the Secretary-General called on the member states, businesses and cross-industry initiatives to advocate for transparent and accountable content governance frameworks that

---

org/ifcn-covid-19-misinformation/, accessed on 22 June 2020.

[237] Based on an interview with Yudhanjaya Wijeratne, Senior Researcher, Data, Algorithms and Policy, LIRNEasia and Co-founder, Watchdog Sri Lanka on 24 June 2020. See more at Yudhanjaya Wijeratne, "Big Tech Is as Monolingual as Americans", *Foreign Policy*, 07 May 2019.

protect freedom of expression, avoid incentives for overly restrictive moderation practices and protect the most vulnerable.[238] But there have always been concerns surrounding content moderation, particularly about the risk of curtailing freedom of speech in the process. In this regard, the UN Strategy and Plan of Action on Hate Speech stresses, "Addressing hate speech does not mean limiting or prohibiting freedom of speech. It means keeping hate speech from escalating into something more dangerous, particularly incitement to discrimination, hostility, and violence, which is prohibited under international law."[239]

The Christchurch Call can be used as a reference to the best practices of international collaboration. The Christchurch shooting in New Zealand in March 2019 which was live streamed, copied and reposted millions of times across several digital media platforms, opened up intense debates regarding the moderation of terrorist and violent extremist content online. Following this, France and New Zealand announced the Christchurch Call, a non-binding document encouraging greater efforts to counter online extremism, apply ethical standards in reporting terrorist events online, and share information between governments and technology companies.[240] A total of 48 countries, three international organizations and leading online service providers have joined the call.[241] As part of the initiative, the online service providers committed to "take transparent and specific measures to prevent the upload of terrorist and violent extremist content and to prevent its dissemination on social media and similar content-sharing services, including its immediate and permanent removal, without prejudice to law enforcement and user appeals requirements, in a manner consistent with human rights and fundamental freedoms."[242] This initiative also led to the re-launching of the Global Internet Forum to Counter Terrorism (GIFCT). The GIFCT was set up in 2017 by Facebook, Microsoft, Twitter and YouTube as an industry-led initiative to apply technology, share knowledge and support research on terrorists' use of platforms. GIFCT will now become an independent entity with dedicated resources and staff and go beyond its initial mission and address both terrorist and violent extremist content online.[243]

---

[238] United Nations, *Report of the Secretary-General: Roadmap for Digital Cooperation*, New York: United Nations, 2020.

[239] United Nations, *UN Strategy and Plan of Action on Hate Speech*, op. cit., p. 1.

[240] Lauren Dudley, "Year in Review: Content Moderation on Social Media Platforms in 2019", available at https://www.cfr.org/blog/year-review-content-moderation-social-media-platforms-2019, accessed on 21 June 2020.

[241] RT Hon Jacinda Ardern, *Significant Progress Made on Eliminating Terrorist Content Online*, Wellington: New Zealand Government, 2019.

[242] The Ministry of Foreign Affairs and Trade, *The Christchurch Call: To Eliminate Terrorist and Violent Extremist Content Online*, Wellington: The Ministry of Foreign Affairs and Trade, New Zealand, 2019.

[243] RT Hon Jacinda Ardern, op. cit.

Additionally, the response of tech companies to the information disorder surrounding the COVID-19 pandemic can be used as another reference. Facebook, Google, Twitter, YouTube, LinkedIn, Reddit, and Microsoft released a joint statement in March 2020 announcing their collaboration in preventing online misinformation and fraud around coronavirus.[244] Based on their individual public policies, the companies on one hand have prioritized official information and on the other, demoted disputed content by decreasing its visibility, providing a warning label or taking it down. While it is argued that misinformation and disinformation still existed in the platforms and flagged contents were not taken down swiftly, this initiative can be seen as a step in the right direction. Such examples of international collaboration and cross-industry initiatives can be used to develop governance frameworks of content moderation related to hate and dangerous speech against race, religion, and ethnic communities and also, hostile information campaigns from foreign actors. South and Southeast Asia need to be active parts of such initiatives owing to the magnitude of information disorder faced by the countries of both regions.

7.11 Engaging in Global Conversations to Make Online Platforms More Accountable

Discussions and debates regarding the accountability and transparency of digital platforms have been on-going for the past few years. The recent surge in the Black Lives Matter movement in the USA brought back the topic in limelight. While the social media platforms were instrumental in mobilizing activists, it was also used for hate speech and misinformation. Different social media companies responded differently to such content. In this regard, the role of Facebook came under severe criticism as the company decided not to act against President Donald Trump's statement which was perceived by many people to have the 'potential to incite vigilante violence'.[245] Such incidents of the company's failure to address disputed content gave rise to protests like 'Stop Hate for Profit' boycott as part of which more than 1,000 companies forswore advertising on Facebook for at least the month of July 2020.[246] The 'Stop Hate For Profit' campaign advocated for establishing and empowering permanent civil rights infrastructure with the expertise to evaluate products and policies for discrimination, bias, and hate.[247] It also called for regularly submitting reports of identity-based hate and misinformation to independent third party audits and making the results publicly accessible.

[244] Kang-Xing Jin, "Keeping People Safe and Informed About the Coronavirus", available at https://about.fb.com/news/2020/06/coronavirus/#joint-statement, accessed on 22 June 2020.

[245] Julia Carrie Wong, "'Too big to fail': why even a historic ad boycott won't change Facebook", *The Guardian*, 11 July 2020.

[246] Ibid.

[247] Stop Hate For Profit, "Recommended Next Steps", available at https://www.stophateforprofit.org/productrecommendations, accessed on 17 July 2020.

Besides this movement, there are several other campaigns initiated in different parts of the world, mostly the USA and Europe, calling on the platforms to make their algorithm transparent and allow people to understand how the filter bubbles and echo chambers shapes their newsfeeds by amplifying a particular line of narrative and cutting down the others. It is important for the countries of South and Southeast Asia to be an active voice in these campaigns and conversations as they are one of the worst victims of the abuse of these platforms.

Chapter 8

Conclusion

In the Information Age, the media landscape evolved in several phases with the advancement of technology as well as social, economic, and political changes. The online platforms of the 21st century gave immense power to people but this power has been abused in different forms. This has been visible in several major events of the world at present. But the scale of the operations and impacts in South and Southeast Asia needs special attention as the countries of these two regions have witnessed some of the worst impacts of abuse of online media platforms and messaging services. Through a contextualized conceptual framework, the paper finds that several actors, such as cyber troops, click farms, fake new syndicates, hard-line religious groups, volunteer groups, paid citizens and individuals have dominated the information ecosystem of these countries with misinformation, disinformation, malinformation, hate speech, defamation, and rumour. The influence of several underlying social, political and economic factors, particularly the religious, ethnic, and racial fault lines play a key role in this regard. Changes in the political environment, media landscape and technological transformation also need to be taken into consideration. Understanding the business model and *modus operandi* of the online platforms is also important as the actors' bank on the technology and features of the companies to amplify their messages and reach audiences on a mass scale and through micro-targeting. The actors have amplified the existing differences in the societies through the adoption of these easy, affordable and hard-to-trace technology. The combination of all these created an information disorder which has serious impacts on the state and society, including vigilante killings, mob attacks and oppression of religious and ethnic minorities. This has also led to xenophobia, damaged social movements and protests, impacted democracies, public institutions, foreign relations and hampered the law and order situation of the country.

Information disorder also has implications for national security. While most of the origins of information disorder in South and Southeast Asia are from domestic sources, hostile information campaign from foreign sources was recently recorded in few countries. This indicates the need for states to focus on this dimension of the problem as well. The influence of international information campaigns was also seen during the global crisis. The 'Infodemic' and war of words among powerful countries surrounding the COVID-19 pandemic led to confusion and uncertainty regarding the pandemic and largely contributed to the information disorder in the hyper-connected virtual world of the two regions.

All these forms and factors of information disorders led many governments to adopt policies and take action. However, there are concerns regarding some of the provisions of the legal frameworks of the countries which need to be reconsidered. Nevertheless, governments across the regions face complex challenges to balance

between security and freedom of speech. Moreover, the tactics involved in the process continuously evolve, circulate among cross-platforms, and often have an online-offline connection which makes it difficult to address. It is also observed that although the problem of information disorder is country-specific, there are transnational and regional implications of the issue as well.

In many cases, it was seen that a particular incident in a country triggers rapid sharing of misinformation, disinformation and hate speech in the neighbouring countries with the same religious and ethnic groups. This can be seen in three forms: similarity in content and narratives among the actors of the same religion, similarity in narrative among actors of different religious communities against a common target group, and support for the victims or target groups by similar religious communities in nearby countries using different forms of disinformation. All these reveal the transnational nature of the problem. There are also regional implications as it is speculated that the availability and affordability of several financially motivated actors in the region can be used to engage in disinformation activities in target countries. Also, there is resemblance among the content shared and its impacts, such as child kidnapping rumours resulting in vigilantism and misinformation surrounding COVID-19. Analysis of the disinformation scenarios in countries of the region with similar socio-cultural contexts can be effective in predicting or assessing potential risks and challenges in one's own country. Additionally, the alarming rise of extremist narratives on online platforms by hardline groups of different religions also have security implications for the region.

The combination of all these factors indicates a complex and quickly evolving challenge for the states that need to be effectively addressed. In this regard, the paper suggests some proactive, reactive, immediate, and long-term approaches that can be adopted by the relevant stakeholders. First, it identifies the initiatives that individual countries can adopt, like promote research and documentation; establish new fact-checking initiatives and facilitate the expansion of existing ones; initiate strong and trusted government communication by creating publicly accessible online platforms that provide clarified version of the information which have potential to impact social cohesion and security; update legal frameworks and approve draft policies, adopt initiatives to safeguard elections, detect foreign hostile information campaigns and take precautionary measures against the spill-over of ongoing tension or conflict along communal lines from other countries of the region. Second, it recognizes the need for a database for curating the disinformation, misinformation, rumours, defamation and hate speech prevalent in the online platforms of the regions for effectively studying the topic and taking preventive measures. It also indicates the possibility of regional cooperation among the BIMSTEC countries in the form of think tank dialogues and Track 1.5 diplomacy for sharing best practices, identifying patterns, and conducting collaborated researches on the topic. Finally, it suggests the need for engaging in the global discussions of content governance, digital cooperation and demand for

making the platforms transparent and more accountable. It is expected that through the collaborated efforts by policy-makers, relevant government agencies, researchers, academics, journalists and fact-checkers, there can be a way forward from the menace of information disorder.

While this paper attempted to provide a comprehensive analysis of the problem and suggest possible ways forward, there are some limitations. There were inadequate academic resources for which the paper could not provide equal emphasis to all countries. Getting access to experts for in-depth interviews from all countries was also not possible within the limited time. However, there are several avenues of the paper which can be considered for further research. For example, the influences of underlying social, political and economic factors of each country can be studied which would give deeper insights regarding the topic. More research can be conducted on the transnational dimensions of the problem to understand how information disorder campaigns in one country influence other countries of the region. Further study can also be carried out on radical narratives and extremist propaganda from different hard-line religious groups in the two regions to assess potential threats and take proactive measures. This paper was an endeavour to shed light on some of the key issues faced within a specific time period but as this is a constantly evolving problem, more research is required. Only the sustained efforts by all stakeholders can bring some order to this intensifying challenge of information disorder.

ANNEX-I

List of interviewees

The following interviewees were consulted during May-July 2020:

- Din M. Sumon Rahman, PhD, Professor, Department of Media Studies & Journalism, University of Liberal Arts Bangladesh
- Yudhanjaya Wijeratne, Senior Researcher, Data, Algorithms & Policy, LIRNEasia and Co-founder, Watchdog Sri Lanka
- Rudroneel Ghosh, Indian analyst and journalist, The Times of India
- Mahbub Roni, Co-founder and Secretary, BD FactCheck
- Saimum Reza Piash, Senior Lecturer at BRAC University and Co-founder at Bangladesh Cyber & Legal Center
- Afia Sultana Pina, Program Coordinator, Promoting Media Literacy in Bangladesh, South Asia Center for Media in Development (SACMID)
- Asish Thakur, Executive Director at Glocal Pvt. Ltd
- Akanksha Narain, Delhi-based independent researcher
- Aisha Nazim, Journalist and Communication Manager, Sri Lanka

The author also received inputs from media practitioners, activists, researchers, regulators, and former employees of social media houses in Bangladesh, India, Sri Lanka, Singapore, and the Philippines who choose to remain anonymous.

ANNEX-II

Check List for Personal Interview

In-depth interviews were conducted through phone conversations and video conferences following a comprehensive check list. However, it was customized based on the profession of the interviewee and his/her country.

1. Please provide an overview of how the information disorder scenario has evolved in your country over the past few years.
2. Who are the major actors behind organized information disorder campaigns in your country?
3. What are the motivating factors behind these actors?
4. Do you think the underlying socio-economic and political factors in your country have a role in the creation and dissemination of such content?
5. Can you please share some of the popular tactics that are used by such actors?
6. What is the extent of disinformation shared through peer to peer messaging applications in comparison to public platforms?
7. Please discuss the role of fact-checkers in your country and elaborate on the scope and challenges of their work.
8. What are the main types of content addressed by fact-checkers?
9. Is there any scope for collaboration with fact-checkers from other countries in the region?
10. What is your opinion on government facilitated/sponsored fact-checking initiatives like in Singapore and Malaysia? Do you think such initiatives could be a feasible solution for other countries of the region?
11. What is your view on the widespread of hate speech and extremist content in social media platforms?
12. Do you think the online platforms are doing enough to contain such kinds of content?
13. What are the existing legal instruments to address disinformation, misinformation, rumour and hate speech on online platforms in your country?

14. Do you think there are definitional challenges of the terms?

15. What are the major impacts of information disorder in your country?

16. Foreign interference through social media is a growing security concern at present, do you foresee any such challenges for your country?

17. Do you think a common pattern can be drawn among the countries of your region regarding information disorder? If so, please share some insights from your work experience.

18. Do you think there are transnational implications of such campaigns?

19. Are there any cross-platform arrangements for detecting disinformation in South and Southeast Asia?

20. Do you think content moderation measures can be taken against hate speech targeting religious and ethnic minority communities in the regions?

21. What will be your suggestions for effectively addressing the challenges related to information disorder?

BIISS Publications

- **Books**

South Asian Regional Cooperation: A Socio–economic Approach to Peace and Stability
Nation Building in Bangladesh: Retrospect and Prospect
The Indian Ocean as a Zone of Peace
The Security of Small States
ASEAN Experiences of Regional and Inter-regional Cooperation: Relevance for SAARC
Development, Politics and Security: Third World Context
Bangladesh and SAARC: Issues, Perspectives and Outlook
Bangladesh: Society Polity and Economy
South Asia's Security: Primacy of Internal Dimension
Chandabaji Versus Entrepreneurship: Youth Force in Bangladesh
Development Cooperation at the Dawn of the Twenty First Century: Bangladesh-German Partnership in Perspective
Conflict Management and Sub-regional Co-operation in ASEAN: Relevance of SAARC
National Security of Bangladesh in the 21st Century
Civil Society and Democracy in Bangladesh
Regional Co-operation in South Asia: New Dimensions and Perspectives
Confidence Building Measures and Security Cooperation in South Asia: Challenges in the New Century
Bangladesh-Southeast Asia Relations: Some Insights
Security in the Twenty First Century: A Bangladesh Perspective
25 Years of BIISS: An Anthology
Politics and Security in South Asia: Salience of Religion and Culture
Small States and Regional Stability in South Asia
Religious Militancy and Security in South Asia
Global War on Terror: Bangladesh Perspective
Towards BIMSTEC-Japan Comprehensive Economic Cooperation: Bangladesh Perspective
Democracy, Governance and Security Reforms: Bangladesh Context
Whither National Security Bangladesh 2007
National Security Bangladesh 2008
Human Security Approach to Counter Extremism in South Asia: Relevance of Japanese Culture
National Security Bangladesh 2009
Energy Security in South Asia Plus: Relevance of Japanese Experience
Changing Global Dynamics: Bangladesh Foreign Policy
Bangladesh in International Peacebuilding: Discourses from Japan and Beyond

South Asia Human Security Series:

Nepali State, Society and Human Security: An infinite Discourse
Evolving Security Discourse in Sri Lanka: From National Security to Human Security
Violence, Terrorism and Human Security in South Asia
Women and Human Security in South Asia: The Cases of Bangladesh and Pakistan
Human Security in India: Health, Shelter and Marginalisation
Pakistan: Haunting Shadows of Human Security
Human Security in India: Discourse, Practices and Policy Implications
Human Security Index for South Asia: Exploring Relevant Issues
Ethnicity and Human Security in Bangladesh and Pakistan

BIISS Publications

- **BIISS Journal (Quarterly)**
- **Bangladesh Foreign Policy Survey (Quarterly)**
- **BIISS Papers (Monograph series)**

The Assam Tangle: Outlook for the Future (1984)
The Crisis in Lebanon: Multi-dimensional Aspects and Outlook for the Future (1985)
India's Policy Fundamentals, Neighbours and Post-Indira Developments (1985)
Strategic Aspects of Indo-Sri Lanka Relations (1986)
Indo-Bangladesh Common Rivers and Water Diplomacy (1986)
Gulf War: The Issues Revisited (1987)
The SAARC in Progress: A Hesitant Course of South Asian Transition (1988)
Post-Brezhnev Soviet Policy Towards the Third World (1988)
Changing Faces of Socialism (1989)
Sino-Indian Quest for Rapprochement: Implications for South Asia (1989)
Intifada: The New Dimension to Palestinian Struggle (1990)
Bangladesh: Towards National Consensus (in Bangla, 1990)
Environmental Challenges to Bangladesh (1991)
The Gulf War and the New World Order: Implication for the Third World (1992)
Challenges of Governance in India: Fundamentals under Threat (1995)
Bangladesh in United Nations Peacekeeping Operations (1998)
Nuclearisation of South Asia: Challenges and Options for Bangladesh (1998)
The Middle East Peace Process and the Palestinian Statehood (2000)
Pakistan and Bangladesh: From Conflict to Cooperation (2003)
Integrated Coastal Zone Management in Bangladesh: A Case for People's Management (2003)
WTO Dispute Settlement System and Developing Countries: A Neorealist Critique (2004)
State Sovereignty and Humanitarian Intervention: Does One Negate the Other? (2006)
Unipolarity and Weak States: The Case of Bangladesh (2009)
Japan's Strategic Rise (2010)
The Fallacy of Fragile States Indices: Is There a 'Fragility Trap'? (2017)
Implications of China's Belt and Road Initiative for Bangladesh: A Strategic Analysis (2020)
Disaster Risk Reduction and Resilience: A Quest for Human Security in Bangladesh (2020)

- **BIISS Seminar Proceedings**

Contemporary Development Debate: Bangladesh in the Global Context
Moving from MDGs to SDGs: Bangladesh Experience and Expectation
SAARC at 30: Achievements, Potentials and Challenges
Bangladesh's Relations with Latin American Countries: Unlocking Potentials
Civil-Military Relations in Democracy: An Effective Framework
Recent Extremist Violence in Bangladesh: Response Options
25 March – Gonohottya Dibosh (Genocide Day)
Reconciling Divided Societies, Building Democracy and Good Governance: Lessons from Sri Lanka
Promoting Cultural Diversity of Small Ethnic Groups in Bangladesh
Upcoming 45th Session of the Council of Foreign Ministers of OIC, Dhaka: Revisiting A Shared Journey
রোহিঙ্গা সংকটঃ বাংলাদেশ কর্তৃক গৃহীত পদক্ষেপ ও পর্যালোচনা (Rohingya Crisis: Measures Taken by Bangladesh and An Appraisal)
Bangladesh Delta Plan 2100
Bangladesh in International Peacebuilding: Experience from Japan
Bangladesh Delta Plan 2100: Implementation, Challenges and Way Forward
1971 Genocide in Bangladesh

BIISS Publications

- **BIISS Seminar Proceedings**

  Bangladesh-India Cooperation: In the Changing Regional and Global Context

  International Day of Peace 2019 and Launching of Book Titled "Bangladesh in International Peacebuilding: Discourses from Japan and Beyond"

  Commemorating the Silver Jubilee of Diplomatic Relation Between South Africa and Bangladesh

  Implications of the Belt and Road Initiative for the Sustainable Development Goals in Bangladesh

  Bangladesh-Nepal Relations: Prospects for Sub-regional Cooperation

  Bangladesh and India: A Promising Future

- **BIISS Country Lecture Series**

  BIISS Country Lecture Series: Part- 1

  BIISS Country Lecture Series: Part- 2