State of Consumer Digital Security in India

Research and Writing by **Pranav M B**
Conceptualisation by **Amber Sinha**
Research Assistance by **Samyukta Prabhu** and **Vibha Nadig**

Introduction

Since 2006, successive Union governments in India have shown increased focus on digital governance. The National e-Governance Plan was launched by the UPA government in 2006, and several digital projects led by the state such as digitisation of the filing of taxes, appointment process for passports, corporate governance, and the Aadhaar programme (India's unique digital identity system that utilises biometric and demographic data) arose under it, in the form of mission mode projects (projects that are part of a broader National e-governance initiative, each focusing on specific e-Governance aspects, like banking, land records, or commercial taxes). In 2014, when the NDA government came to power, the National e-Governance Plan was subsumed under the government's flagship project of Digital India, and several mission mode projects were added. In the meantime, the internet connectivity, first in the form of wire connectivity, and later in the form of mobile connectivity has increased greatly. In the same period, use of digital services, first in new services native to the Internet such as email, social networking, instant messaging, and later the platformization and disruption of traditional business models in transportation, healthcare, finance and virtually every sector, has led to a deluge of digital private service providers in India.

Currently, India has 500 million internet users — over a third of its total population — making it the country with the second largest number of Internet users after China. The uptake of these technological services has also been accompanied by several kinds of digital threats that an average digital consumer in India must regularly contend with. This report is a mapping of consumer-facing digital threats in India and is intended to aid stakeholders in identifying and addressing digital security problems. The first part of the report categorises digital threats into four kinds, Personal Data Threats, Online Content Related Threats, Financial Threats, and Online Sexual Harassment Threats. Threats under each category are then defined, with detailed consumer-facing consequences, and past instances where harm has been caused because of these threats.

Personal Data-related Threats

With the rapid rise in the availability of personal information that exists online, there is a huge threat to consumer privacy online. This section will analyse the range of digital threats related to the collection, processing, sharing and retention of personal data. The use of personal data for profiling and personalised targeting of content and advertising has emerged as a major vector of manipulation online. This has happened primarily over social media and online messaging services.

1. Doxing

Doxing (or doxxing) is the act of publishing or broadcasting private or identifying information (often personally identifiable information[1]) about a person or a group of persons. The broadcast or publishing happens on the Internet, with platforms ranging from popular social media websites to semi-private online discussion forums.[2]

Impact

Publishing someone's personal information online, especially without their consent can lead to several harmful consequences that range from damage to reputation, harassment/bullying, and even financial theft. The magnitude of harm depends on the nature, and sensitivity of information that is published.[3]

A prominent case of doxing in India took place in 2018, where a journalist and author was subjected to severe online harassment after personal details about her, including her address, were shared online.[4] More recently the practice of doxing has gained much more attention in the context of the Hong Kong protests, where protesters, journalists, police officers and trolls have had their personal information revealed online without consent.[5]

2. State Surveillance

Surveillance is the monitoring of users' behaviour, activities, and information, carried out by government bodies, often with the involvement of private sector parties.

Impact

While considered essential for intelligence gathering, and maintaining law and order, unrestricted surveillance can be excessively pervasive and pose a threat to users privacy. Additionally, when surveillance is carried out without adhering to secure data protection practices, the user data collected through surveillance runs the risk of suffering a breach

---

[1] Personally identifiable information refers to information that can be used either on its own or in combination with other information to identify, locate or contact an individual.
[2] C.S-W. (2014, March 10). What doxxing is, and why it matters. Retrieved from https://www.economist.com/the-economist-explains/2014/03/10/what-doxxing-is-and-why-it-matters.
[3] Liebl, L. (2014, October 28). The dangers and ramifications of doxxing and swatting. Retrieved from https://www.gamezone.com/originals/the-dangers-and-ramifications-of-doxxing-and-swatting/.
[4] The Wire Staff. (2018, April 28). Delhi Journalists Body Condemns Relentless Trolling of Rana Ayyub. Retrieved from https://thewire.in/media/delhi-journalists-body-condemns-relentless-trolling-of-rana-ayyub.
[5] Lim, L. (2019, November 11). Doxxing: the powerful 'weapon' in the Hong Kong protests had a petty beginning. Retrieved from https://www.scmp.com/magazines/post-magazine/short-reads/article/3036663/doxxing-powerful-weapon-hong-kong-protests-had.

and posing a privacy threat. Additionally, lack of effective regulation to restrict the usage of collected data could lead to the data being transferred to third parties for their utilisation, again creating additional threats to personal data.

In India specifically, the minimal restriction on government surveillance has been flagged as a major data privacy concern in the 2019 Forrester Global Map of Privacy.[6] In 2018, when the government's ID database suffered a breach, a large number of private actors began selling this leaked personal data. This breach was also considered the largest breach of data globally, according to the World Economic Forum's 2019 Global Risks Report.[7]

Compounding with the limited legislative oversight there is on India's surveillance and intelligence gathering regime, there are numerous cases where intelligence bodies have been given more power to conduct surveillance. The National Cyber Coordination Centre is an operational e-surveillance agency set up to scan and record metadata on the Internet.[8] In 2011, the Research and Analysis Wing was authorised to intercept phone calls, emails and voice and data communications domestically.[9] Presently, ten bodies in the country derive legislative and executive authority to monitor all computer resources.[10]

3. Vulnerabilities in Online Architecture

Online services that fail to meet adequate security standards are vulnerable to attacks from malicious actors. These actors can take advantage of flaws or loopholes in the online architecture to gain unauthorised access to personal data. Both public and private entities that collect data are vulnerable in these cases, especially if they do not follow the required security practices.

Impact

Similar to the consequences that are seen in some cases of surveillance, the breach of personal information is a threat to privacy, and can lead to a range of harmful outcomes ranging from harassment, and monetary loss, and can also transition into other threats

---

[6] Sheth, H. (2019, August 1). 'Government surveillance at alarming levels'. Retrieved from https://www.thehindubusinessline.com/info-tech/government-surveillance-at-alarming-levels/article28138407.ece.
[7] Sapkale, Y. (2019, February 19). Aadhaar Data Breach Largest in the World, Says WEF's Global Risk Report and Avast. Retrieved from https://www.moneylife.in/article/aadhaar-data-breach-largest-in-the-world-says-wefs-global-risk-report-and-avast/56384.html.
[8] Salman, S.H. (2017, August 10). National Cyber Coordination Centre: Govt's online metadata scanning project is live. Retrieved from https://www.medianama.com/2017/08/223-national-cyber-coordination-centre-launch/.
[9] TNN. (2011, December 19). RAW gets power to tap phones, track emails. Retrieved from https://timesofindia.indiatimes.com/india/RAW-gets-power-to-tap-phones-track-emails/articleshow/11161977.cms.
[10] Singh, N. (2018, December 21). The Indian government puts all computers under surveillance. Retrieved from https://www.medianama.com/2018/12/223-all-computers-india-surveillance/.

like doxing. If there is unauthorised access to a lot of personal data, this data could also be utilised to analyse behaviour, activity, and preferences, and indirectly monitor people.

Due to a security breach by Truecaller, about 299 million Indian mobile numbers leaked, along with other information such as people's email IDs and subscriber photos, as of May 2019.[11] Several other private entities have faced instances of this nature where either an accidental breach or a malicious attack results in a large threat to personal data. Pune-based instant digital-lending startup EarlySalary was the victim of a ransomware attack in October 2018. The attacker was looking to extort ransom against data of at least 20,000 users that they had accessed from an earlier version of its website.[12] In January 2019, Amazon India also admitted to a technical glitch that left many vendors' data exposed.[13]

4. Data Leaks Caused Due to Human Negligence

A non-technical vulnerability is often the human element involved in the collection and processing of data. If a data leak occurs due to human error, or intentional malice, similar consequences follow as they would if a technical vulnerability was exploited.

In 2017, sensitive data belonging to at least ten banking companies was accidentally leaked on the popular code sharing and version control service, Github. The leak was attributed to a developer working at Tata Consultancy Services (TCS),[14] and while no consumer data was leaked, such leaks are so commonplace that Fallible, a Bangalore-based cybersecurity firm, built Gitleaks.com, a tool that scanned terabytes of public data on Github for patterns of exposed secrets, such as database credentials, passwords, and private keys.[15] Records of malicious data theft/leaks also are commonplace, often sharing personal client data to corporate rivals for monetary gain.[16]

---

[11] Sarmah, H. (2019, May 22). Data Breach: Truecaller Exposes Indian Users' Data, Shows Cracks In Cyber Security Infrastructure. Retrieved from https://analyticsindiamag.com/data-breach-truecaller-exposes-indian-users-data-shows-cracks-in-cyber-security-infrastructure/.

[12] Bhakta, P. (2018, October 5). EarlySalary says data of 20k stolen. Retrieved from https://economictimes.indiatimes.com/small-biz/startups/newsbuzz/earlysalary-says-data-of-20k-stolen/articleshow/66079892.cms?from=mdr.

[13] Ghosh, S. (2019, January 10). Data breach at Amazon India exposes sellers' financial data. Retrieved from https://yourstory.com/2019/01/amazon-india-sellers-data-leak-breach.

[14] Sharma, S. (2017, June 14). TCS employee accidentally leaks confidential data on Github, gets roasted. Retrieved from https://factordaily.com/tcs-employee-github-data-leak-jason-coulls/.

[15] Fallible. (2017, February 19). We are resuming GitLeaks.com. Retrieved from https://hackernoon.com/we-are-resuming-gitleaks-com-d93cf73824e3.

[16] TNN. (2018, December 1). Ex-employee leaks data to rival firm, held. Retrieved from https://timesofindia.indiatimes.com/city/noida/ex-employee-leaks-data-to-rival-firm-held/articleshow/66889320.cms.

5. Lack of Adequate Breach Disclosure and/or Response Measures

Exploited vulnerabilities result in incidents where personal data is leaked. Subsequently, the private or public entity that suffered the breach is tasked with improving their security to protect against future attacks, as well as taking measures to minimise damage caused by the present attack.[17] This usually involves disclosing information about the attack to either the public, affected parties, or a regulatory entity. When adequate information is not disclosed, or is disclosed after a delayed period, it reduces the opportunities of affected parties to independently take precautionary measures to safeguard their personal data. If entities that collect data do not have, or do not practice adequately effective breach disclosure measures, the potential data threat that exists in each breach is higher because of lack of capacity to safeguard data subsequent to the breach.

The food delivery service provider FreshMenu suffered a data breach in 2016, exposing the personal data of a large number of users, but this breach only became public knowledge in 2018.[18] In January 2019, while Amazon India disclosed that a data breach had occurred, it refused to divulge the extent of the breach and its nature.

6. Collection of unnecessary personal data and data retention

If there is a collection of more data than necessary, in case of a potential data leak/security breach, there is an increased threat to people's privacy, due to the chance of more personal data being leaked to unauthorised parties. Furthermore, in cases where necessary or unnecessary data is collected but not properly disposed of, the same risk exists where more personal data could be leaked to unauthorised parties.

Data collection entities like Truecaller,[19] and Facebook[20] are known to collect large amounts of data which do not prima facie appear to aid in the services the entities

[17] Malik, Y. (2019, January 10). Another data breach? Amazon India leaks sellers information in tech error. Retrieved from https://www.business-standard.com/article/companies/another-data-breach-amazon-india-leaks-sellers-information-in-tech-error-119011000001_1.html.
[18] Goswami, S. (2018, September 14). FreshMenu Hid Data Breach Affecting 110,000 Users. Retrieved from https://www.bankinfosecurity.asia/freshmenu-hid-data-breach-affecting-110000-users-a-11514.
[19] Krishnan, R. (2019, May 22). Real threat: Truecaller data available for sale. Retrieved from https://economictimes.indiatimes.com/tech/internet/real-threat-truecaller-data-available-for-sale/articleshow/69437379.cms.
[20] Whittaker, Z. (2019, June 13). Facebook collected device data on 187,000 users using banned snooping app. Retrieved from https://techcrunch.com/2019/06/12/facebook-project-atlas-research-apple-banned/.

provide. Furthermore, Facebook's data retention policies indicate that some personal data will be retained indefinitely, and may never be deleted.[21]

7. "Privacy Policy" related Threats

Entities that transact with personal data regulate their interactions with users and their data through a Privacy Policy. Threats to personal data can arise if the entities fail to:

a. Make the privacy policy available
b. Draft a policy that complies with required data protection regulations
c. Communicate the policy, and obtaining educated consent

The failure of making a policy available to users creates a lack of accountability on the part of the company collecting the data, to its users. Since there is no document outlining their promise to their users regarding protecting their information, they can get away with potential security breaches/invasion of privacy through collection of information (which could further be shared with third parties).

The failure to draft a policy that complies with the required data protection regulations also shows lack of responsibility on the part of the company to its users, who would assume that the company would comply with the legally drafted data protection regulations. This threat continues to remain in India without adequate data protection regulations in place.

Apart from communicating the policy to its users, the company should actively obtain educated consent from its users, i.e. ask explicit permission from users before collecting, storing, processing, sharing, or undertaking any other activity with their data. If there are pre-checked boxes and extremely long privacy policy documents, it is highly unlikely that users will provide educated consent to the companies for usage of their data.

An evaluation of the privacy policy of consumer applications in India showed that privacy policies are generic, unavailable in Indian languages, and offered no examples of layered notices.[22] For instance, Flipkart's privacy policy, which runs for about 1,800 words, fails to mention data retention or deletion policies. Similarly, Paytm's 642 word privacy policy mentions that user data will not be shared with third parties for unsolicited marketing. But the term "third parties" is undefined. Furthermore, privacy policies of smaller Indian start-ups seem to put the onus completely on users regarding updates to the privacy

[21] Picchi, A. (2018, March 23). OK, you've deleted Facebook, but is your data still out there?. Retrieved from
https://www.cbsnews.com/news/ok-youve-deleted-facebook-but-is-your-data-still-out-there/.
[22] Christopher, N., Bansal, V. (2018, June 18). Indian startups still don't take data privacy seriously. Srikrishna Committee may change it soon. Retrieved from
https://economictimes.indiatimes.com/small-biz/startups/newsbuzz/indian-startups-still-dont-take-data-privacy-seriously-srikrishna-committee-may-change-it-soon/articleshow/64628026.cms.

policies. Users are required to check for updates to the privacy statement in case of changes and there is no voluntary notification (In contrast, larger entities like Facebook, Google, and Whatsapp notify users of changes in their privacy policies).

8. Lack of Security of Data Transmission Channels

When communication involving transfer of personal data between two systems is intercepted by an outside entity, there is a clear personal data threat. These attacks, popularly referred to as “Man in the Middle” attacks (MITM) usually happen when third parties take advantage of free and insecure wifi services, perform email hijacking, or perform website session hijacking.[23] Instances of attacks of this nature have led to significant harms including leak of personal data as well as financial loss.[24]

Online Content-related Threats

Digital media, especially social media and online messaging apps, provide a new and viral mode of circulation of information. While these new mediums were intended to play an important role in increasing awareness, it has also resulted in much greater dissemination of fake news and misinformation. Online media is also ripe with instances of other forms of problematic content, most notably hate speech relating to caste and religion, and online harassment of users online, including doxing. Some of these instances have also had very direct repercussions in the offline world, including lynchings and physical threats. It would also be relevant to study which groups and demographics are most vulnerable to the influence of these propaganda campaigns.

In the process of mapping online content related threats, one key link that has been identified is the chilling effect of online harassment. When acts of bullying, trolling, abuse, and other forms of harassment happen online, they result in what is essentially silencing of different voices online that are either targeted by the harassers, or are observing the harassment and do not want to undergo similar experiences. In both cases, parties often choose to remain silent to avoid becoming targets of harassment. Therefore, for the purposes of this mapping, sources of harassment are considered online threats.

[23] Publico, R. (2017, March 1). What is a Man-in-the-Middle Attack and How Can You Prevent It?. Retrieved from https://www.globalsign.com/en-in/blog/what-is-a-man-in-the-middle-attack/

[24] Naidu, J.S. (2017, January 23). Man-in-the-middle case: Mumbai firm loses Rs10.89 lakh to online fraudster. Retrieved from https://www.hindustantimes.com/mumbai-news/man-in-the-middle-case-mumbai-firm-loses-rs10-89-lakh-to-online-fraudster/story-xp3AcjLmnh0vAeY8rUIWYO.html.

1. Fake News

Fake news refers to false information that is published as if it were truthful, and then spread via news media or social media. While the term “fake news” has received criticism for being excessively reductive, the more nuanced categories of this online content related threat are:[25]

a. **Disinformation** False information that is deliberately created, and published with the intent of causing harm to a person, social group, organisation, or country. This includes manipulated and fabricated content that is published as news with the intent of causing harm.
b. **Misinformation** False information that is published, but not created with the intention of causing harm. This includes publishing of false and misleading content often with the belief that the content is true, without the existence of an intent to fabricate or mislead.

A third category “Malinformation”, while not strictly dealing with falsified information, does cause similar kinds of harm. Malinformation refers to truthful or genuine information that is published with the intention of inflicting harm on a person, social group, organisation or country. This can also happen by moving to the public space information that is designed to remain private.[26]

In India, the consequences of disinformation efforts were especially prevalent during the 2019 Indian general election. Political parties extensively utilised social media and communication platforms like Facebook and Whatsapp, publishing numerous pieces of falsified or misleading content to attempt to influence the electorate, ahead of the election.[27] The nature of falsified content across social media, television, and newspapers ranged from criminal allegations, to falsified public statements, from false claims of international pre-election surveys to digitally altered images depicting falsified records. A lot of the false content produced and circulated was divisive in nature, based on religious, regional, caste, and economic differences.[28]

[25] UNESCO. Journalism, 'Fake News' and Disinformation: A Handbook for Journalism Education and Training. Retrieved from https://en.unesco.org/fightfakenews.
[26] Wardle, C., Derakhshan, H. (2017, September 27). Information Disorder: Toward an interdisciplinary framework for research and policy making. Retrieved from https://rm.coe.int/information-disorder-report-november-2017/1680764666.
[27] Sidharth, A. (2019, May 18). How misinformation was weaponized in 2019 Lok Sabha election – A compilation. Retrieved from https://www.altnews.in/how-misinformation-was-weaponized-in-2019-lok-sabha-election-a-compilation/.
[28] Gilbert, D. (2019, April 11). Modi's trolls are ready to wreak havoc on India's marathon election. Retrieved from https://www.vice.com/en_us/article/597mwk/modis-trolls-are-ready-to-wreak-havoc-on-indias-marathon-election.

In November 2019, a study by EU DisinfoLab reported that as many as 265 fake local news websites in more than 65 countries, including the US, Canada, Brussels and Geneva, were managed by Indian influence networks with the aim of influencing international institutions along with elected representatives and swaying the international community's perception of Pakistan.[29]

2. Doxing as a Threat to Free Speech

The act of doxing, in addition to being a personal data threat, can also be used to threaten online free speech. This is done when an entity online threatens to commit the act of doxing, and uses this threat to restrict another user's free speech. Among reported instances, journalists especially have been targeted by this specific threat.[30]

3. Threats arising out of Community Standards

Social media platforms and discussion forums usually are bound by a set of "community standards" which are the equivalents of rules or policies that regulate behaviour on the platforms and lay down restrictions on acceptable conduct. There are two types of online content related threats that can arise out of community standards.

The first is when the standards do not effectively regulate online behaviour/conduct and therefore leave behind unaddressed loopholes that serve as opportunities for harassment.[31] The second type of threat that can exist is connected to the enforcement of community standards. Even though an online entity has adequate safeguards in its rules or standards, there have been numerous instances where the entity does not actively enforce them or does so incorrectly, or non-uniformly. This can exist in the form of removal of posts that do not have problematic content, or selectively removing some of the posts that violate standards instead of applying standards uniformly. [32]

If these entities selectively remove certain types of posts, their labelling of that content as problematic is not based on uniform grounds. This is an online content related threat since it amounts to censoring of information relating to some topics, which could involve

---

[29] Bhargava, Y. (2019, November 14). 265 fake news websites in over 65 countries managed by Indian influence networks: study. Retrieved from https://www.thehindu.com/news/national/265-fake-news-websites-in-over-65-countries-managed-by-indian-influence-networks-study/article29967820.ece.
[30] Reporters Committee staff. (2015, May 19). The dangers of doxxing. Retrieved from https://www.rcfp.org/journals/news-media-and-law-spring-2015/dangers-doxxing/.
[31] Dixit, P. (2017, October 17). Twitter Has A Harassment Problem In India, And Targets Say The Company Isn't Doing Much To Fix It. Retrieved from https://www.buzzfeednews.com/article/pranavdixit/twitter-india-harassment-problem.
[32] Salim, M. (2019, June 13). Facebook's Uneven Enforcement of Hate Speech Rules in India Highlighted in New Study. Retrieved from https://thewire.in/media/facebook-hate-speech-guidelines-india-study.

either raising awareness about sensitive topics (eg. war crimes in Syria[33]), or could be about a topic that is politically charged (eg. removal of LGBT content on YouTube[34]).

Financial Threats

1. Identity Theft

Identity theft is the crime of using someone's personal information, credit history or other identifying characteristics in order to make purchases or borrow money without that person's permission. In India identity theft accounts for 77% of the fraud cases in 2015.[35]

Impact

The impact on the first tier is to the individual whose identity is appropriated. Purchases and borrowing can impact an individuals' credit, as well as their ability to engage in future transactions. It can also cause security breaches and pin liability on the victim of identity theft.

The second tier impact is regarding the market created for identities. A high demand has created a market for identities to be sold on the dark web. The result has been an increase in verification and securities in order to discern real identities. This has caused friction between consumers and corporations.[36]

SIM fraud is one particular manifestation of identity theft. This is contingent on the perpetrators' ability to port the number to a new SIM. When this is done, the old SIM is cancelled, locking the victim out. Having the SIM allows access to the two step authentication. Victims in India have lost more than Rs. 200 crore via SIM fraud in 2018.[37]

---

[33] Rosen, A. (2018, March 7). Erasing History: YouTube's Deletion Of Syria War Videos Concerns Human Rights Groups. Retrieved from https://www.fastcompany.com/40540411/erasing-history-youtubes-deletion-of-syria-war-videos-concerns-human-rights-groups.
[34] Bardo, S. (2018, January 17). YouTube Continues To Restrict LGBTQ Content. Retrieved from https://www.huffpost.com/entry/youtube-continues-to-restrict-lgbtq-content_b_5a5e6628e4b03ed177016e90.
[35] Securitas. (2017). Identity theft is the largest contributor to fraud in India. Retrieved from https://www.securitas.in/globalassets/india/files/about-us/news---related-documents/identity-theft-is-the-largest-contributor-to-fraud-in-india.pdf.
[36] Nicolls, D. (2018, November). Cybersecurity Threats Facing Financial Services. Retrieved from https://www.jumio.com/5-cybersecurity-threats-financial-services.
[37] TOI. (2019, January 4). What is a SIM swap fraud? What are the safety tips?. Retrieved from https://timesofindia.indiatimes.com/business/india-business/what-is-a-sim-swap-fraud-what-are-the-safety-tips/articleshow/67377708.cms.

2. Social Engineering

Social engineering refers to a fraudulent technique that involves obtaining sensitive information from people, often by communication falsified content. This can happen over the phone, email and even in-person.[38]Today, only about 3 percent of malware relies solely on bugs or technical glitches. The other 97 percent targets its users through social engineering. Nearly 60 percent of security leaders say their organizations may have fallen victim to social engineering within just the past 12 months.[39]

Impact

Social engineering is designed to enable the perpetrator to convince the victim to divulge information related to their bank account, credit card, etc. With reference to online payments through UPI, fraudsters often pose as customer care, requesting OTP or personal details. The result is often loss of huge amounts of money in the guise of completing a bank transaction, or availing an exclusive offer. Some examples of such social engineering fraudulent activities are:

a. **Phishing** A malicious party sends fraudulent communication disguised as an authentic source.
b. **Watering Hole** Injecting malware into public websites.
c. **Whaling Attack** An advanced form of phishing characterized by the groups it targets- government officials and bureaucrats.
d. **Baiting and Quid Pro Quo Attacks** Offering an advantage or a good in exchange for a particular action.
e. **Tailgating** Gaining access to restricted information by “piggybacking” off of those who have been authenticated.

3. Technical Vulnerabilities

A technical vulnerability is an error, flaw, failure or fault in a computer program or system that causes it to produce an incorrect or unexpected result, or to behave in unintended ways.

Possibly the biggest financial frauds in recent years occurred when Rs 25 was moved out of Bank of Maharashtra (BoM) accounts due to a bug in the UPI application. Another case was when media reports came out showing that the Truecaller app had sent out SMS messages

[38] Nicolls, D. (2018, November). Cybersecurity Threats Facing Financial Services. Retrieved from https://www.jumio.com/5-cybersecurity-threats-financial-services.
[39] Ibid.

from the phones of unsuspecting users to create UPI IDs with ICICI Bank without their consent. These bugs are widespread and cause harms on an individual level as well.[40]

Impact

Vulnerabilities in software can enable technology to be used in harmful ways. Often times, these are exploited by fraudsters in order to enable their own enrichment. The biggest problem with bugs in platforms such as BHIM and UPI is not just the danger of the bug itself, but the difficulty in pinning liability on either the platform or the bank.

4. Account Takeover

Account takeover is using another person's account information (e.g., a credit card number) to obtain products and services using that person's existing accounts. In 2018, India reported about 1.4 lakh account takeover (ATO) login attempts every hour from people using stolen, or generated usernames and passwords. According to the 2018 'Credential Stuffing: Attacks and Economies' report by global firm Akamai, India was the second most preferred target destination after the US, recording more than 120.8 crore ATOs in just the one year.[41]

Impact

ATO-based attacks are particularly dangerous and effective because they often originate from accounts of trusted senders via phishing attacks. This has two significant consequences: First, the ATO is probably going to succeed on the grounds that there is a previous trust association with the client. Second, these assaults frequently go undetected by customary security controls since they start from genuine records.

One particular form of Account Takeover that impacts users via BHIM, UPI is SIM swapping. SIM swapping refers to a type of account takeover fraud which bypasses the protection provided by two-factor authentication. This method is effective when the user has lost their phone, or is switching to a new phone. At this point, through social engineering techniques, the attacker can convince the telephone service provider to port the victim's phone number to the attacker's SIM card. This leads to the cancellation of the old SIM card, subsequently resulting in all two factor verification (OTP) functions reaching the attacker instead of the user, rendering the protective tool useless.

[40] Vijaysarathy, S. (2019, August) UPI Bug and Truecaller. Retrieved from https://www.indiatoday.in/technology/news/story/truecaller-upi-bug-is-this-spam-filter-app-reading-your-bank-sms-to-create-your-credit-profile-1575998-2019-08-01.
[41] Kumar, C., (2019, August). Cyber Threat in India. Retrieved from https://timesofindia.indiatimes.com/business/india-business/cyber-threat-india-witnessed-1-4-lakh-account-hacking-attempts-every-hour-in-2018/articleshow/69019739.cms.

5. Ransomware

Ransomware is a type of software that holds data and computers 'hostage' by restricting access, until a certain amount of money is paid as ransom. This is often triggered by phishing, most commonly via email. In 2017, financial services were the second most targeted industry of ransomware after healthcare.

Impact

The threat is not just to financial institutions who must bear the cost, but also the data of individuals who consume the services of the institutions. Privacy is at the forefront of this threat, as a non-payment of the amount would lead to the release of sensitive and private information into the public domain.[42]

Regulatory Frameworks

1. Financial Threats

a. Payment Regulations

The Payment and Settlement Systems Act, 2007 ('PSS Act, 2007'), legislated in December 2007, provides for the regulation and supervision of payment systems in India. Under the PSS Act, 2007 the RBI is given the power to direct and regulate the payment systems and the payment system participants in India. This act created the Payment and Settlement Systems Regulations, which regulates consumer-relevant matters like the determination of standards of payment systems, furnishing of returns/documents/other information, and furnishing of accounts and balance sheets by system providers.

b. Technology regulations

The Information Technology Act, 2000 provides legal recognition to the transaction done via electronic exchange of data and other electronic means of communication or electronic commerce transactions. This also involves the use of alternatives to a paper-based method of communication and information storage to facilitate the electronic filing of documents with the Government agencies.

[42] Kaspersky Labs. (2018, August). Ransomware and Malicious Crypto. Retrieved from https://securelist.com/ransomware-and-malicious-crypto-miners-in-2016-2018/86238.

Specifically Section 66C covers identity theft, and Section 66D covers cheating by personation by using computer resources.

c. **Banking Regulations**

It was issued in order to provide a framework for the regulation and supervision of persons operating payment systems involved in the issuance of Pre-paid Payment Instruments (PPIs) in the country and to ensure the development of this segment of the payment and settlement systems in a prudent and customer friendly manner.

The terms are applicable to all transactions initiated by the User through Bharat Interface for Money application(“**BHIM App**”) developed by National Payments Corporation of India (“**NPCI**”) and using Unified Payments Interface (“**UPI**”) services as a mode for the transfer of funds. Notably, NPCI make no warranties about the quality of service, and disclaim all liability arising out of harm caused to consumers.[43]

2. Online Content Threats

a. **Facebook**

Community Standards outlined by Facebook mention five areas of regulation.

i. Violence and criminal behaviour
ii. Safety
iii. Objectionable content
iv. Integrity and authenticity
v. Respecting intellectual property

Anything that is interpreted by Facebook to be in violation of the above standards can be taken down from the platform. Compliance to standards is often regulated by AI. Facebook attempts to combat fake news through broadly three ways:

---

[43] The disclaimer states, “NPCI does not hold out any warranty and makes no representation about the quality of the UPI Services or BHIM application. The User agrees and acknowledges that NPCI shall not be liable and shall in no way be held responsible for any damages whatsoever whether such damages are direct, indirect, incidental or consequential and irrespective of whether any claim is based on loss of revenue, interruption of business, transaction carried out by the User, information provided or disclosed by Issuer Bank regarding User’s Account(s) or any loss of any character or nature whatsoever and whether sustained by the User or by any other person. While NPCI shall endeavour to promptly execute and process the transactions as instructed to be made by the User, NPCI shall not be responsible for any interruptions, non-response or delay in responding due to any reason whatsoever, including due to failure of operational systems or any requirement of law.”

i. disrupting economic incentives because most false news is financially motivated,
ii. building new products to curb the spread of false news; and
iii. helping people make more informed decisions when they encounter false news.

News Integrity Initiative[44]

Facebook has created a group with participants including tech industry leaders, academic institutions, non-profits and third party organizations — to launch the News Integrity Initiative, a global consortium focused on helping people make informed judgments about the news they read and share online.

Facebook Journalism Project[45]

Facebook is convening key experts and organizations doing important work in this area, such as the Walter Cronkite School of Journalism and Mass Communication at Arizona State University, and have been listening and learning to help decide what new research to conduct and projects to fund. Working with the News Literacy Project, they are producing a series of public service announcements (PSAs) to help inform people on Facebook about this important issue.

b. WhatsApp

Even though WhatsApp is primarily a messaging services, with the kind of unprecedented use that it sees in India, it has also become the most platform most affected digital threats like fake news and hate speech. Some of the regulatory measure undertaken by WhatsApp are:

i. Fake News Tip Line

WhatsApp unveiled its "Checkpoint Tipline," where people can check the authenticity of information received as the messaging giant looks to crack down on fake news, launched by PROTO and WhatsApp. Once a WhatsApp user shares a suspicious message with the tipline, PROTO's verification

[44] Brown, C. (2017, April 2). Introducing the News Integrity Initiative. Retrieved from https://www.facebook.com/facebookmedia/blog/introducing-the-news-integrity-initiative.
[45] Simo, F. (2017, January 11). Introducing the Facebook Journalism Project. Retrieved from https://www.facebook.com/journalismproject/introducing-facebook-journalism-project.

centre will seek to respond and inform the user if the claim made in the message shared is verified or not.[46]

ii. **Forward Limit & Label**

The app shows a forwarded label on the top left corner of the messages. While this feature was introduced as a cautionary tale in July last year, later it restricted the forwarding limit to five chats to hinder the flow of such fake content.[47]

iii. **Developmental Ideas**

Image Search Feature WhatsApp is reportedly working on an image search feature that will enable users to verify the authenticity of an image by directly uploading the image on to Google, following which it will open the browser for you to see the search result.[48]

In-app Browser Another feature that WhatsApp is hoping to use to combat fake news is its in-app browser. It would help in informing users if their requested page is not safe to visit and if the site they requested has fake content.[49]

c. Quora

The platform does not have explicit mechanisms to combat fake news. However, content is regulated via their Privacy Policy and Acceptable Use Policy. The AUP has generic provisions regarding no online bullying, hate speech, and intellectual property rights.

Under its FAQs, there are answers pertaining to spam and "factually incorrect content". On Quora, spam is defined as one or more questions, answers, posts, comments, or messages whose purpose appears to be to direct traffic to external commercial sites while providing little to no value back to the Quora Community. If

[46] Ghoshal, A. (2019, March). WhatsApp launches a tip line in India to battle fake news ahead of national elections. Retrieved from https://thenextweb.com/apps/2019/04/02/whatsapp-launches-a-tip-line-in-india-to-battle-fake-news-ahead-of-national-elections/.
[47] Hern, A., Safi, M. (2019, January 21). WhatsApp puts limit on message forwarding to fight fake news. Retrieved from https://www.theguardian.com/technology/2019/jan/21/whatsapp-limits-message-forwarding-fight-fake-news.
[48] Carman, A. (2019, March 18). WhatsApp tests in-app reverse image searches to prevent the spread of hoaxes. Retrieved from https://www.theverge.com/2019/3/18/18270890/whatsapp-browser-reverse-image-search-beta.
[49] Ibid.

repeated "factually incorrect content" is posted, it is treated as spam, post which the post is collapsed.[50]

Astroturfing is when a user creates a large amount of content on Quora at a fast pace with a desire to promote content that is deceptive or factually incorrect. Astroturfing is also considered spam on Quora and may also violate the policies on self-plagiarism.[51]

d. Twitter

Twitter challenges millions of potentially spammy accounts every month — in 2018, on average between 8.5 million and 10 million accounts — by requesting additional authentication details (like email addresses and phone numbers). Between January and June of 2018, about 75% of the accounts challenged were suspended.

Election Precautions: During the 2019 Lok Sabha elections in India, Twitter set up a separate category for reporting misleading information regarding the elections.[52] There were three categories of content that Twitter sought to take down:

i. Misleading information about how to vote or register to vote (for example, that you can vote by Tweet, text message, email, or phone call);

ii. Misleading information about requirements for voting, including identification requirements; and

iii. Misleading statements or information about the official, announced date or time of an election.

e. Legally mandated regulation

There are three sections that pertain to online content threats under the Information Technology Act, 2000

i. **66A.** Punishment for sending offensive messages through communication service, etc.

ii. **67A.** Punishment for publishing or transmitting of material containing sexually explicit act, etc., in electronic form.

iii. **67B.** Punishment for publishing or transmitting of material depicting children in sexually explicit act, etc., in electronic form

[50] Quora. (2019, January 13). How does Quora define spam? What are the consequences of spamming on Quora?. Retrieved from https://help.quora.com/hc/en-us/articles/360000470266-How-does-Quora-define-spam-What-are-the-consequences-of-spamming-on-Quora-.

[51] Ibid.

[52] Lomas, N. (2019, April 24). Twitter to offer report option for misleading election tweets. Retrieved from https://techcrunch.com/2019/04/24/twitter-to-offer-report-option-for-misleading-election-tweets.

Additionally, the Information Technology [Intermediaries Guidelines (Amendment) Rules] 2018 categorically specifies that the intermediaries must inform to the users of the computer resource about the Rules and regulations and privacy policy so as to not to host, display, upload, modify, publish, transmit, update or share any information which might affect public health and safety and Critical Information structure.

The 2018 Rules further provides that whenever an order is issued by the government agencies seeking information or assistance concerning cyber security, then the intermediaries must provide them the same within 72 hours.

3. Online Sexual Harassment Threats

a. Criminal Law Provisions

Sections 354A and 354D of Indian Penal Code provide punishment for cyber bullying and cyber stalking against women.

Section 509 of IPC comes to your rescue if someone is constantly bugging you with derogatory verbal abuse because of your gender. The section provides that any person who utters any word or makes any sound or gesture, intending that such word, sound or gesture be heard or seen by a woman and insult her modesty, shall be punished with one-year imprisonment and/or fine.

b. Regulations for Protection of Children

The Ministry of Women and Child Development had enacted the Protection of Children from Sexual Offences Act, 2012 (POCSO Act) as a special law to protect children from offences of sexual assault, sexual harassment and pornography. Section 13 to Section 15 deals with the issue of child pornography.

Section 14 and Section 15 lays down the punishment for using children for pornographic purposes and for storage of pornographic material involving children.

c. Regulations for Technological Offences

Transmission or publication of material containing sexually explicit acts is punishable under the Information Technology Act. There is also a separate provision dealing with the publishing or transmitting of material depicting children in sexually explicit act, etc., in electronic form.

d. Workplace Harassment Regulations

Online harassment also encompasses sexual harassment which is defined under section 2(n) of the Sexual Harassment of Women at Workplace (Prevention, Prohibition, and Redressal) Act 2013 as unwelcome conduct such as a demand or request for sexual favours; or making sexually coloured remarks; or showing pornography. Behavior conducted over an digital medium including social networking channels, where it would fall within the domain of workplaces are covered under this legislation.

4. Personal Data Regulation

a. Personal Data Protection Bill, 2019

The Bill defines personal data as data about or relating to a natural person who is directly or indirectly identifiable.

The Bill allows data processing by fiduciaries if consent is provided by the individual. However, in certain circumstances, processing of data may be permitted without the consent of the individual. These include (i) any function of Parliament or state legislature, or if required by the State for providing benefits to the individual, (ii) if required under law or for compliance with any court judgement, (iii) to respond to a medical emergency, or a breakdown of public order, (iv) purposes related to employment, such as recruitment, or, (v) for reasonable purposes specified by the Data Protection Authority with regard to activities such as fraud detection, debt recovery, credit scoring, and whistle blowing.

The Bill sets out certain rights of the data principal whose data is being processed. These include (i) the right to obtain a summary of their personal data held with the data fiduciary, (ii) the right to seek correction of inaccurate, incomplete, or outdated personal data, (iii) the right to have personal data transferred to any other data fiduciary in certain circumstances, and (iv) the right 'to be forgotten', which allows the data principal to restrict or prevent continuing disclosure of their personal data.

b. The Aadhaar and Other Laws (Amendment) Bill

The Act provides for the use of Aadhaar number as proof of identity of a person, subject to authentication. The Bill replaces this provision to state that an individual may voluntarily use his Aadhaar number to establish his identity, by authentication or offline verification. The Bill states that authentication of an individual's identity via Aadhaar, for the provision of any service, may be made mandatory only by a law of Parliament.

Under the Act, restrictions on security and confidentiality of Aadhaar related information do not apply in case the disclosure is pursuant to an order of a District Court (or above). The Bill amends this to allow such disclosure only for orders by High Courts (or above). Further, under the Act, an officer not below the rank of a Joint Secretary may issue directions for disclosing information in the interest of national security. The Bill amends this to allow such disclosure on directions of officers not below the rank of a Secretary.

c. Information Technology Act

Under section 69 of the IT Act, any person, authorised by the Government or any of its officer specially authorised by the Government, if satisfied that it is necessary or expedient so to do in the interest of sovereignty or integrity of India, defence of India, security of the State, friendly relations with foreign States or public order or for preventing incitement to the commission of any cognizable offence relating to above or for investigation of any offence, for reasons to be recorded in writing, by order, can direct any agency of the Government to intercept, monitor or decrypt or cause to be intercepted or monitored or decrypted any information generated, transmitted, received or stored in any computer resource.

d. Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules

The rules provide the reasonable security practices and procedures, which the body corporate or any person who on behalf of body corporate collects, receives, possess, store, deals or handle information is required to follow while dealing with "Personal sensitive data or information". In case of any breach, the body corporate or any other person acting on behalf of body corporate, the body corporate may be held liable to pay damages to the person so affected.

e. DNA Technology (Use and Application) Regulation Bill

The DNA Bill seeks to create a national DNA data bank, and regional data banks, and have separate indexes for crime records such as crime scene, undertrials, suspects, missing persons and unknown deceased persons. It has proposed DNA sampling and profiling of citizens accused of crime or reported missing, and storing their unique genetic information for administrative purposes. Under the proposed DNA Bill, the government seeks to harvest the data with consent from some people, and without consent from others, as in the case of individuals who have committed serious crimes attracting imprisonment beyond seven years.

The Road Ahead

On 26 July 2018, Ravi Shankar Prasad, the IT Minister of India was giving a speech in the Rajya Sabha, the upper house of the Indian Parliament. He warned that social media platforms cannot 'evade their responsibility, accountability and larger commitment to ensure that its platform is not misused on a large scale to spread incorrect facts projected as news and designed to instigate people to commit crime.' More ominously, he said that if 'they do not take adequate and prompt action, then the law of abetment also applies to them.' The Minister was speaking in response to the rising incidents of mob lynchings in India, ostensibly occasioned by the spreading of misinformation, inciting violence on social media and mobile messaging services. Comparing social media services to newspapers, Prasad further said that when there is provocative writing in newspapers, the newspaper cannot say that it is not responsible.

Prasad's words of warnings were not in isolation. Since the revelation about Cambridge Analytica's use of Facebook to profile and manipulate users with political content emerged, the Indian government has been engaged in a series of ad hoc communications with large Internet intermediaries.

When the Cambridge Analytica-Facebook data incident was widely reported, the Ministry for Electronics and IT sought details from both Cambridge Analytica and Facebook about how many Indian residents' data was impacted during the incident. While Cambridge Analytica did not provide a clear response, Facebook admitted that the data of 560,000 Indians were compromised.

One of the key problems with the intermediary liability regime is that it paints all intermediaries with the same brush. As mentioned above, the term applies equally to internet service providers, social media platforms, search engines, private messaging services, e-commerce companies and web-hosting services. The new proposed rules also do not disturb this status quo. The basis for safe harbour is the idea that intermediaries are mere dumb conduits for the distribution of the speech of its users, rather than speakers themselves. However, this argument of the dumb conduit is no longer tenable. Most, if not all intermediaries affirmatively shape the form and substance of user content in some manner, using highly intelligent prioritization algorithms. A multi-pronged approach that involves different kinds of regulatory tools that cut across the range of digital threats we face, is sorely needed.