Traceability in End-to-End Encrypted Environments

Technical and Socio-Economic Impacts of Implementing Traceability in India

12 August 2024

Context

In February 2021, the Government of India introduced the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules (or IT Rules 2021) under the parent legislation, the Information Technology Act, 2000 (or IT Act 2000). The primary aim of the IT Rules 2021 is to regulate Internet services and intermediaries, including messaging services, and news media intermediaries, like digital media houses.

The IT Rules 2021 impose new conditions on intermediary liability protections—or safe harbor—in the Indian context. Section 79 of the IT Act 2000 exempts intermediaries from liability for third-party content on the intermediaries' platforms, as long as platforms comply with the IT Act. These legal protections from liability for user-generated content have enabled people across the country to participate on the Internet—creating, contributing, and communicating content. However, the IT Rules 2021 impose new compliance requirements on intermediaries, limiting their protection from liability if they are not met.

One of these new compliance requirements is 'traceability': a requirement of significant social media intermediaries providing services primarily in the nature of messaging. Significant social media intermediaries include platforms with over 5 million registered users. Traceability requires platforms offering end-to-end encrypted messaging to identify the first originator of messages shared on their service. This will include services such as (but not limited to) Signal, WhatsApp, and iMessage. The first originator is assumed to mean the first person who shares a message on the platform. If that person is based outside the country, the first person to share the message within India's territory will be considered the first originator.

Several requirements in the IT Rules 2021 have been challenged through at least 20 independent lawsuits filed by various organizations and individuals. In at least three cases challenging the legality of traceability, the petitioners contend that it deprives people of the freedoms guaranteed under the Constitution of India and breaks the essence of end-to-end encryption, thus undermining the security and privacy of encrypted communication. The High Court of Delhi is now hearing these petitions.

This brief aims to inform the legal and public discourse on traceability, by explaining how it impacts the security and privacy of end-to-end encrypted communications for everyone—including children, government, and businesses. Further, we highlight the technical flaws and security vulnerabilities that traceability would introduce.

What Is End-to-End Encryption?

End-to-end encryption (E2EE) delivers the highest technical level of confidentiality and is the strongest type of encryption. This technology ensures that nobody apart from the sender and receiver of information has access to the content, not even the service provider. The sender is one 'end' of the conversation, and the receiver is another 'end'—ensuring no third party can access it. Strong encryption, especially E2EE, keeps all of us safe online and offline.

Proposed Technical Approaches to Implement Traceability on End-to-End Encrypted Services

While not specified under the IT Rules 2021, the government has proposed two primary ways in which traceability could be implemented:

The Kamakoti proposal: The identity of the originator of each message will be tagged and included in an encrypted form with the message. Formulated by Professor V. Kamakoti of the Indian Institute of Technology, Madras, this proposal[1] will require an intermediary to hold the key to decrypting the originator's information in escrow (for the government). In case of an order from an authorized body, the intermediary will then be required to share this information with the government. This proposal also suggests that users mark a message 'forwardable' or 'non-forwardable' to indicate whether they intend the message to be forwarded. If a user originates a message and marks it as 'forwardable,' their information gets linked with the message. However, if a sender marks a message 'non-forwardable' and the recipient forwards it nonetheless, the recipient becomes the first originator, and their information, in turn, is linked to the message.

Hashing: This method proposes that messaging services create alpha-numeric 'hashes' of all the messages exchanged over their platform before being encrypted on the end-user device. Hashing is the practice of using an algorithm to map data to a fixed-length numeric or alpha-numeric string. For instance, if someone sends "Hello" over an E2EE messaging service, it may generate a (hypothetical)

[1] Kamakoti, V. (2019) *Report on Originator Traceability in WhatsApp Messages*. Available at https://dn790005.ca.archive.org/0/items/reportofprof.kamakotiinwpnos.20214and20774of2018/Report%20of%20Prof.%20Kamakoti%20in%20WP%20Nos.20214%20and%2020774%20of%202018.pdf (Accessed 12 August 2024)

hash of "R9T," but "hello"—with one lowercase character "h"—will generate a completely different hash. These hashes will form a library of hashes to be retained on the platform's servers, against which the hash of a message can be compared to enable traceability, subject to a court or government order. An E2EE messaging service would, therefore, have to maintain a library of hashes to allow the government to trace the originator of a message in case of an authorized order.

Traceability Undermines Encryption and Our Safety

Encryption enables us to keep our personal information safe. Whether we share family pictures with loved ones or our children's location when picking them up, E2EE ensures that our sensitive information is secure from bad actors. E2EE messaging apps are used as a substitute for in-person, face-to-face conversations. The nature of the data shared on E2EE platforms is thus usually highly confidential and sensitive.

E2EE also offers marginalized, at-risk groups the safety of communicating anonymously or pseudonymously. Any traceability requirement is based on a presumption that anonymous or pseudonymous communication is not legitimate, which is a dangerous premise.

Criminals and malicious actors could gain access to sensitive and personal information if encryption is weakened. Forcing platforms to weaken security offered by strong encryption by implementing traceability would be detrimental to the safety, security, privacy, and livelihood of people, businesses, and governments worldwide. It would also result in severe financial losses[2]: first, by undermining security and making people vulnerable to cyber-attacks and criminals, and second, due to an erosion of trust in secure, private communications. Weakening encryption also undermines confidence in investors, as they may not view India as a safe place to invest or do business.

E2EE systems are designed to provide communication security, confidentiality, and integrity. They are not intended to identify a message's originator or trace how it is shared. Traceability requirements would require a complete redesign of a service provider's infrastructure and compromise the technical measures that ensure the confidentiality and integrity of encrypted communications. Implementing traceability in E2EE environments would introduce a deliberate, systemic vulnerability in messaging systems.

Regarding the two methods proposed by the Indian government, cybersecurity experts have identified these concerns[3]:

---

[2] Internet Society (2021) *The Economic Impact of Laws that Weaken Encryption.* Available at https://www.internetsociety.org/resources/doc/2021/the-economic-impact-of-laws-that-weaken-encryption/ (Accessed 12 August 2024)

[3] Internet Society (2020) *Traceability and Cybersecurity: Experts' Workshop Series on Encryption in India.* Available at https://www.internetsociety.org/resources/doc/2020/traceability-and-cybersecurity-experts-workshop-series-on-encryption-in-india/ (Accessed 12 August 2024)

Kamakoti Proposal

Embedding originator information in a message will need the encryption technology to be altered. It would introduce backdoors or weaknesses in the security offered by E2EE, opening up sensitive personal information to potential attacks by bad actors. Cybercriminals could intercept and manipulate originator information, leading to false attribution of messages. Simply put, innocent people threaten to be falsely caught as tech-savvy criminals may use modified apps to potentially frame others.

Digital attribution, as proposed by Prof Kamakoti, is not absolute and is vulnerable to impersonation. The use of a person's device is not necessarily evidence of the person's use of the device. To establish criminal liability, guilt has to be proven beyond reasonable doubt—a threshold that is hard to meet, especially when impersonation online is so easy and pervasive. Cybercriminals can easily impersonate someone else's digital ID, thus falsely implicating innocent people.

Private keys held by messaging platforms will be valuable targets for bad actors. This could lead to unauthorized access and tampering by service provider employees and contractors and external attacks that could breach the service provider's infrastructure, leading to mass-scale cyberattacks and data breaches.

Such a proposal also conflicts with the principle of data minimization. The principle of E2EE holds that the moment any information can be tracked or indeed 'traced' by authorities—in this instance, by messaging platforms holding originator information—it breaks E2EE.

In countries like India, where disinformation has—in the past—led to the loss of life and property, such vulnerabilities could give rise to more viral, harmful rumors.

Hashing

Generating hashes for each message also undermines confidentiality. When a hash is generated for a message on an E2EE messenger, it takes into account the unique identity keys of the sender and receiver in question, in addition to the content of the message. This means that the hash of "Hello" sent from Roy to Zeba will be different from the "Hello" sent or forwarded from Zeba to Sam. If Sam were to report this message to the authorities, and messaging platforms could search their hash libraries, only the hash generated by Zeba to Sam would show up—not the hash generated by Roy to Zeba. Even two messages from Roy to Zeba saying "Hello" would have different hashes.

Hashing will require messaging platforms to track every message sent on their service, defeating the purpose of E2EE. By having access to and gathering so much information, the proposal of hashing also undermines data minimization.

Hashes can also be reverse-engineered, especially for common phrases. Stored in large libraries on the messaging platforms' servers, this could also undermine the confidentiality of messages. For instance, anyone could calculate the hash of a message that contains commonly used phrases. If they had access

to the library of hashes as proposed by this method, they could identify everyone who has sent that exact message. Also, since hashing takes place on an end-user device, this opens up opportunities for manipulation by bad actors.

Even other proposals that seek access to information on E2EE messaging platforms while claiming to maintain the integrity of encryption, suffer from drawbacks. In a few jurisdictions, there has been a proposal to use metadata to enable traceability.

Metadata

Digital attribution through metadata—simply, data describing other data—lacks certainty and is not absolute. Minor changes to a message can alter its metadata. Using metadata for developing social graphs to enable traceability poses significant risks, as criminals can compromise these graphs.

Not all platforms collect the same type or even amount of metadata. Some services are designed to minimize metadata collection as a legitimate privacy-preserving measure, and they should not be required to collect more data simply to comply with the mandate of traceability. Reliance on metadata undermines message confidentiality by contravening data minimization and privacy by design principles, which are increasingly mandated by data protection policies.

The use of metadata has been proposed in other countries but has not been accepted. In Brazil, for instance, the use of metadata was previously proposed to contain virality, i.e., restrict the spread of viral messages. However, doing so based on metadata alone—with no knowledge of the content of the message—is dangerous. Messages can go viral for legitimate reasons, such as warnings about an approaching natural disaster like a storm or tsunami.

Metadata retention rules often require extended retention periods, creating a honeypot of information that could be a target for cybercriminals.

In summary, traceability requirements:

- Undermine the security and confidentiality of end-to-end encrypted communications; and
- Expose and collect information about any person who forwarded a message, increasing the risk of data breaches and misuse of that information by criminals for extortion, phishing, and other attacks.

The proposals also fail to consider the variety of message-forwarding methods that are available to users beyond forwarding a message within a service, including copy and paste within or between messaging services, re-entering a message, sending screenshots rather than text, and using extra-territorial sources. These failures mean that the traceability measures are likely to be ineffective and fail necessity and proportionality tests required for constitutionality, while still increasing the cost and complexity of critical online services.

These proposals do not account for human behavior and error. A slight mistake in typing, the accidental pressing of an alphabet key, mistakenly forwarding a message to an unintended recipient, or devices slipping from people's grip while they are texting, can lead to severe and disproportionate consequences.

Traceability Also Presents Socio-Economic Challenges

There is widespread consensus among cybersecurity and technical experts, public policy advocates, and civil society that traceability cannot systemically address the problem of getting to people who deliberately and consistently spread malicious information, fake news, and rumors. Instead, it will disproportionately harm innocent people.

Traceability threatens the security and privacy of everyone using E2EE messaging services. This is because the implementation of traceability will require services to trace non-Indian users too. Thus there is no way for traceability to be implemented in India without damaging the technology for all users worldwide.

Further, it is impossible to ascertain without doubt that the first originator of a message is indeed within the territory of India without adding pervasive and intrusive geo-location tracking. Even then, user accounts which are being used in other countries but on an Indian phone number can create attribution errors. Users could also mask their real location using VPNs and other technologies to avoid being identified as the first originator.

There is an added consideration that a significant section of the population in India struggles with digital literacy and will perhaps find it hard to comprehend the changes in terms of use which will be necessitated by the implementation of traceability. They may unknowingly use messaging apps to share sensitive or confidential data without understanding the associated risks.

The national economy can also suffer as businesses face increased costs associated with data breaches, fraud, and reduced consumer trust. Strong encryption is crucial for securing financial information. Just as businesses and banks employ armored truck services to protect money transfers in the physical world, online financial services rely on encryption to safeguard their clients' finances and transactions. Robust security standards and protocols are essential for boosting user adoption of online financial services like online banking and digital payments. When the pandemic hit, India already had the highest volume of digital transactions worldwide, and this number is expected to increase five-fold by 2025. This growth has been made possible by security protocols to protect transactional data. To sustain this growth, user confidence must remain intact. Weakened encryption standards resulting from the implementation of traceability obligations may severely undermine such confidence.

Implementing traceability may potentially create new vulnerabilities, adversely affecting healthcare service providers, as they use E2EE messaging for consultations and sharing of medical information with

patients. The introduction of traceability could jeopardize patient confidentiality and the secure exchange of information, potentially leading to privacy breaches and misuse of sensitive health data.

India's digital infrastructure, while improving, still faces challenges related to cybersecurity and may be ill-equipped to handle these additional risks. Traceability mechanisms may lead to under-preparedness for new cyber-attacks, operational challenges, and increased costs for maintaining compliance with new regulations.

People worldwide have invested significant resources—both human hours and financial commitments—to enable seamless, global, secure communication. By introducing a deliberate, systemic vulnerability in E2EE communication, traceability makes people across the world less safe, as it is not possible for the technology to have a built-in vulnerability in one jurisdiction, while being completely secure in another. This way, traceability threatens to undermine decades' worth of connectivity and security efforts for all people.

Finally, end-to-end encryption as a technology is publicly available. Anyone can build an app or service which is secured by E2EE. If traceability is implemented on large messaging platforms, criminals and bad actors could simply move to another platform, or even build another platform for themselves. It is legitimate users who would be left without the vital confidentiality currently offered by popular messaging services.

Conclusion

This analysis has shown that there is no neat, proportionate way to implement traceability—or indeed any backdoor or scanning mechanism—on E2EE systems. Traceability requirements introduce complexities into secure communication services, making them more vulnerable to attack, and adding cost and operational burdens for service providers. The technical feasibility of traceability mechanisms is untested and entirely open to question, particularly at the scale required for global messaging services. Traceability introduces cost, complexity, and systemic vulnerability for an unproven, speculative outcome.

Traceability also threatens fundamental and constitutional rights, including the right to freedom of speech and expression under Article 19(1)(a) and the right to privacy under Article 21 of the Constitution of India. Along with weakening individual security and privacy, traceability threatens to infringe upon the constitutionally upheld right to secure communications for trade and profession under Articles 19(1)(g) and 21 of the Constitution. Such a mandate is neither necessary, reasonable, nor proportionate, nor does it meet the standard of being the least restrictive means of achieving the government's objective for proposing traceability.

In addition to its impact on citizens' rights, the imposition of traceability will also have economic impacts on India as a place to invest and do business. Millions of citizens and businesses rely on

confidential communication to transact, to trade across borders, and to innovate. Loss of confidentiality in those communications undermines trust and makes India a less attractive place to invest and do business.

In summary, traceability impacts security, confidentiality, cost, citizens' rights, commercial trust, economics, and investment—and may prove ineffective when deployed at scale. The usual tests used to mitigate such impacts are those of necessity, proportionality, and effectiveness: we believe the traceability proposals fail those tests.

The Internet Society urges the Government of India to withdraw this requirement and lend support to robust standards for security, confidentiality, and privacy—all of which are afforded to us by strong end-to-end encryption.

This brief has benefited from legal inputs provided by Apar Gupta, Advocate – Law Office of Apar Gupta and Bharucha and Partners. This brief does not constitute legal advice.