RESEARCH REPORT

Key Findings: A Mechanism for Market Monitoring of Customer-facing Issues with Unified Payments Interface (UPI)

Authors

Abhishek Mukherjee

Dwijaraj Bhattacharya

July 2026

Acknowledgements

The authors of this report would like to thank all the members of Dvara Research for their continuous input and feedback on the various aspects of this project. In particular, the authors thank Dr. Indradeep Ghosh (Ex- Executive Director, Dvara Research) and Ms. Deepti George (Ex-Deputy Executive Director & Head of Strategy, Dvara Research) for their continued guidance and support.

As part of the project, the authors have sought and received advice from various experts. The authors thank Dr. Daryl Collins (Chief Executive Officer, Decodis), Mr. Pravarakya Reddy Battula (Director – Analytics, Decodis), Mr. Madhusudhanan S (Group Chief Technology Officer, Dvara Holdings) and Mr. Kundan Joshi (Senior Software Architect) for their valuable time in reviewing the project and providing feedback.

We would also like to acknowledge the efforts by former team members who contributed during various stages of the projects. These include, Anubhutie S, Rakshith P, Shreya S, and Anukriti T.

List of Tables and Figures

Section-1: Introduction

The Unified Payments Interface (UPI) represents one of India's most significant financial technology innovations. Developed by the National Payments Corporation of India (NPCI) under the regulatory oversight of the Reserve Bank of India (RBI), UPI is a real-time payment system that allows instant inter-bank transfers via mobile apps. The NPCI describes UPI as “a system that powers multiple bank accounts into a single mobile application (of any participating bank), merging several banking features, seamless fund routing, and merchant payments into a single hood”.

The UPI system operates through a five-party model for real-time transaction processing:

- Party-1: Payer Payment Service Provider (PSP): The mobile application or bank interface used by the sender to initiate payments, e.g., PhonePe, Google Pay, or bank-specific apps.
- Payee Payment Service Provider: The application or banking interface where the recipient receives funds, which may be the same or different from the payer's PSP.
- Remitter Bank: The payer's bank responsible for debiting the sender’s account.
- Beneficiary Bank: The payee's bank that receives and credits funds to the recipient's account.
- UPI Switch (by NPCI): The central switching infrastructure that facilitates secure communication between all participating banks and validates transaction authenticity.

According to the recent NPCI data[1], UPI processed over 20 billion transactions worth INR 24.85 lakh crore (approximately $298 billion) in August 2025. Year-on-Year, UPI recorded a robust 34 percent increase in transaction volume and 21 percent growth in transaction value[2]. As of 2025, UPI serves over 491 million individual users and 65 million merchants across a network of 675 connected banks[3]. It also commands an 85 percent share of all digital transactions in India and powers nearly 50 percent of global real-time digital payments. Market leaders, like PhonePe and Google Pay, maintain dominant positions with 45.74 percent and 35.30 percent market shares respectively as of August 2025.

The simplicity of making payments and the widespread acceptance among merchants ensure that customers across income levels rely heavily on UPI as their primary transaction method. Given this popularity, even a 0.1 percent transaction failure rate can impact nearly 6.45 million transactions in a single day and 20 million transactions in a month. In reality, the top ten UPI remitter banks reported transaction failures ranging from 0.01 to 0.75 percent in August 2025[4]. Anecdotal evidence suggests that most transaction failures do not result in money being deducted from the remitter’s account, and when they do, transactions are reversed almost immediately. However, there are cases where money is deducted, and the customer cannot receive support from the banks or PSPs regarding their transaction. Customers also face other hurdles, such as accessing the PSP’s platform, server downtime, and fraudulent actors. The data on these hurdles often reside with the PSPs alone and are not visible to regulators. The RBI Ombudsman may have visibility over some of the hurdles but given that only a small and highly selective group of dissatisfied users escalate issues to this level, the quantum is expected to be negligible. This is because filing a formal complaint requires time, awareness, and some understanding of the redressal process. In many cases, customers simply choose not to pursue the matter further. Consequently, Ombudsman data reflects only the remaining

[1] Sourced from the NPCI website (as of September, 2025)
[2] As reported by the Economic Times (September, 2025)
[3] Press Information Bureau (July, 2025)
[4] Sourced from the NPCI website (as of September, 2025)

disputes, which under-represents both the frequency and severity of the problems experienced by the broader customer base.

Therefore, there is a need to develop a mechanism that can help both PSPs and regulators identify and classify issues faced by UPI users in near-real-time. To this end, we developed a tool that can identify customer-facing issues in the UPI ecosystem by monitoring publicly available social media posts. The tool gathers data from X (earlier "Twitter") and Google Play Store[5] for four PSPs (having combined market share of over 90 percent) of transaction volumes. The tool then employs Natural Language Processing (NLP) and Machine Learning (ML) techniques to categorise customer hurdles, estimate their prevalence rates, and produce other useful indicators.

Some key findings from this effort are:

- The prevalence of customers facing UPI-related issues is episodic and event-driven rather than constant. Spikes in complaint volumes and shifts in dominant issue types are closely associated with periods of outages, technical changes, or rapid user onboarding phases. This underscores the value of near-real-time market monitoring as an early warning system, enabling PSPs and regulators to detect emerging risks before they crystallise into large-scale customer harm.
- The typology of customer issues is concentrated around a small set of inter-related concerns. Across both Twitter and the Play Store, transaction failures, technical issues, and inadequate or delayed redress consistently emerge as the most prevalent concerns. Furthermore, these issues frequently co-occur, suggesting that customer dissatisfaction often stems from cascading failures wherein, for example, an initial technical or transaction-related problem is compounded by ineffective grievance redress mechanisms.
- Different platforms surface different dimensions of customer distress. Twitter functions primarily as a channel for escalation with an emphasis on redress-related failures and transaction breakdowns, while Play Store reviews capture a wider spectrum of issues, particularly onboarding and access-related problems. This divergence indicates that relying on a single data source would systematically understate certain categories of customer grievances.

This report presents a brief synopsis of the methodology (Section-2) and anonymised findings (Section-3). A detailed discussion of the latest trends, PSP-level information, and the insights that can be derived is avoided given the limited scope of this document and ongoing provider and policy-level engagements. Specifically, Section-3 is not intended as a showcase of findings; rather, it is intended as a demonstration of the kinds of information that a social media monitoring tool can provide. Finally, additional insights derived from sentiment analysis, response time analysis, LLM/agentic-AI integration, etc. are avoided, given these methodologies are being fine-tuned.

[5] Though Play Store is not technically a social media platform, given the reviews may indeed function as such, the term is used throughout the document for simplicity.

Section-2: Methodology

To operationalise the market monitoring solution, three broad steps are needed. First, identifying the range and prevalence of hurdles faced by customers; second, developing a typology of these hurdles, or customer issues for consistent classification; third, automating the classification process. The following segment describes the methodology used for each of the steps.

1. Data Collection and Issue Discovery: For this study, data has been continuously gathered from four PSPs: Google Pay, PhonePe, Paytm, and NPCI BHIM, using X (earlier and interchangeably referred as "Twitter" throughout the report) and the Google Play Store (hereafter referred to as Play Store) from 1 January 2020. These platforms were chosen because the first three PSPs account for over 90 percent of transactions, and BHIM was chosen, given it was developed by the NPCI itself. Further, a high level of user activity was also found on Twitter and Play Store. These PSPs were also actively engaging on these platforms.

After data was collected from the API and other sources, it was cleaned using standard practices in machine learning and data science. For Twitter, this involved removing duplicates, handle names, non-English words, emoticons, special characters, and all tweets with fewer than two words to maximize coherent information for analysis. The Play Store data was cleaned similarly but with some differences, such as removing responses by PSPs and duplicate reviews. Reviews with more than four stars were ignored for the purposes of user hurdle identification, as these were more likely to highlight positive performance rather than customer issues.

To better understand the range of issues faced by customers using UPI and assess their prevalence, Gibbs Sampling Dirichlet Multinomial Mixture (GSDMM), an unsupervised machine learning algorithm, was run on the initial batch of data collected and cleaned until June 2022. The resulting document clusters were manually tagged to identify specific issue types. Both the GSDMM cluster analysis and manual tagging of the documents were performed through multiple iterations to establish a final typology of customer-facing issues (listed in the table below).

Table 1: Typology of Customer Issues in UPI

For Twitter:	For Play Store:
Access Issues	Access Issues
Technical Issues	Technical Issues
Inadequate/Delayed Redress	Inadequate/Delayed Redress
Phishing	Onboarding Issues
Transaction Failure	Transaction Failure
Third-party Issues	Third-party Issues

We do not describe the specific nature of each issue type because these are mostly self-explanatory categories while also being broad. For example, "Phishing" is self-explanatory, but the content of the tweets varies in terms of the source, method, and other characteristics. Further, it should be noted that the customer issues in this typology are not mutually exclusive. This is an important consideration for classification algorithms in machine learning and should be kept in mind when training models. This is discussed as part of the next step in the methodology.

This typology forms the basis for classifying all data from both Twitter and Play Store. However, considering the third objective, which was to automate the classification, it was necessary to train a model that can be used to identify and classify collected data uniformly and consistently in near-real time.

2. Training models: Manually classified data from January 2020 to June 2022 was used to train a set of models based on the Convolutional Neural Network (CNN) architecture. The reasons for selecting the CNN architecture were primarily motivated by two factors, its ability to capture local n-gram features and computational efficiency, allowing for rapid retraining in the event that accuracy diminishes over time. In the end, multiple CNN models, each with a binary classifier (one for each of the issue types) were selected. Forward testing of the models revealed accuracies ranging from 87 percent to 94 percent.

3. Automating the process: The final step was to automate the entire data collection, cleaning, and classification processes into a single pipeline that allows for meaningful visualisation. To achieve this, the following architecture was set up using Amazon Web Services (AWS):

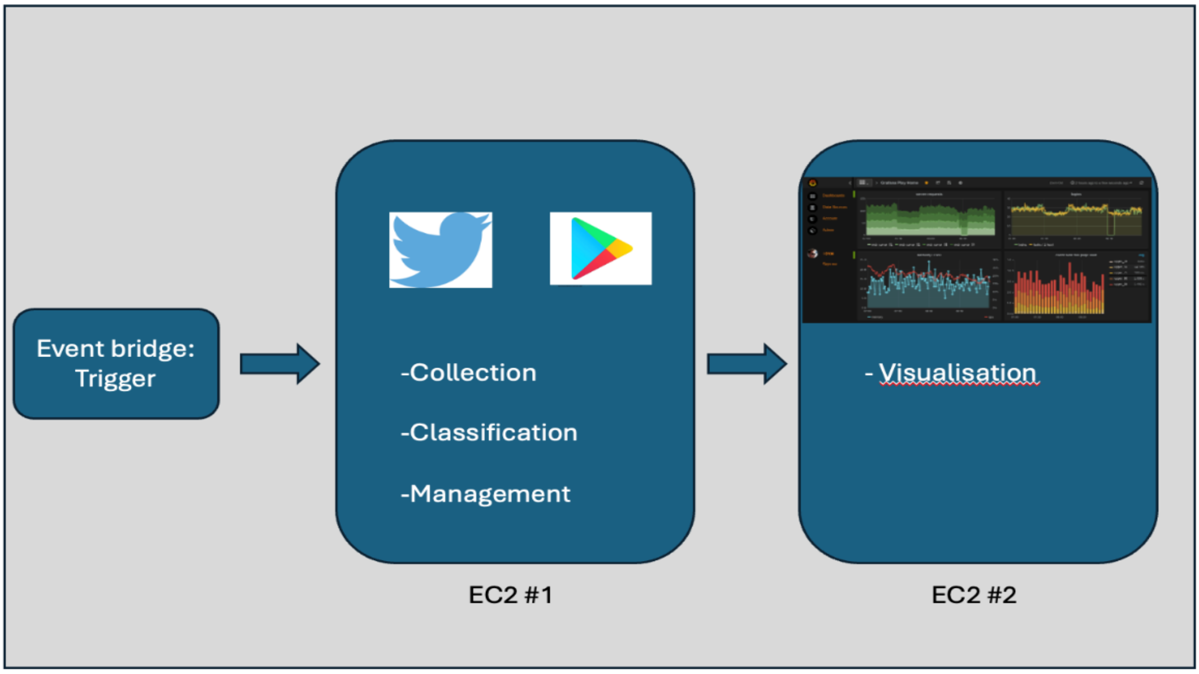


Figure 1: Architecture for automation of the pipeline

The automated pipeline can be broadly understood as consisting of three steps, as presented in Figure 1. The first is the process initiation mechanism, achieved by creating an EventBridge rule that triggers Lambda functions on a predefined schedule. These Lambda functions perform the collection, classification, and management of data. The final step is the visualisation of the data. A detailed description of this process is provided below.

As shown in Figure 1, two instances of Elastic Compute Cloud (EC2) were set up on AWS: EC2 #1 for collecting and processing data, and EC2 #2 for visualising the data. Inside EC2 #1, there are two Docker 'containers' representing the two social media platforms (Twitter and Play Store), each comprising three parts:

- Data collection layer: In-house tools developed by the Dvara Research team for collecting and pre-processing the necessary data from Twitter & Play Store.
- Classification layer: The previously described CNN classification models were used to classify the data according to the developed typology of customer issues.
- Data management layer: Consisting of databases, log storage, etc. for storing the collected data in a standardised format.

EC2 #2 follows EC2 #1, i.e., EC2 #2 takes the classified data and translates it into different visualisations that describe the daily, monthly, and yearly volume of customer issues, the prevalence of such issues across different PSPs, the percentage of classified issues out of the total number of collected tweets/reviews, co-occurrence of these issues, etc. This was done using Grafana, an open-source interactive visualisation web application.

Section-3: Findings

The results from the exercise are presented below in two parts. The first part discusses the results from Twitter and the second part from Play Store. As discussed earlier in Section-1, the data presented is almost a year old and has been anonymised. These actions are deliberate, because of ongoing engagements. Additionally, in most cases the implications are not discussed, this is done since several of the visualisations are either deliberately limited in their temporal coverage or must be read alongside others for accurate inference, given that each can be explained through multiple causal pathways. Discussing each of these hypotheses and their validation would detract significantly from the objective of this document, and as such, is not presented.

Part 1: Twitter

Figure 2 shows the monthly volume of tweets from January 2020 to June 2025. In a few cases PSPs had multiple handles, and as such, the results presented are an aggregate of all of these handles. While we can present the daily volume of tweets (refer Appendix 1.1) for the entire study period, here we discuss only the monthly volume for easier visualisation. Based on daily volume counts, Provider 3 had the highest number of tweets, with a daily average of 340.13 during the entire study period. This was followed by Provider 4, with an average of 97.08 tweets per day; Provider 2 at 66.62 and Provider 1 at 29.78 tweets.

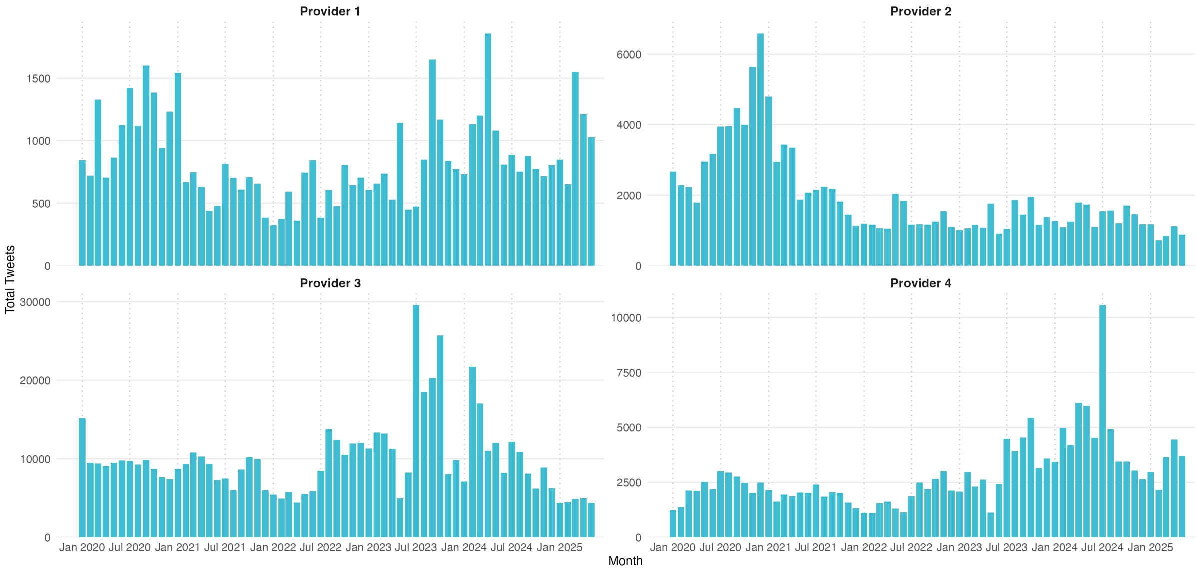


Figure 2: Monthly Volume of Twitter

Based on monthly volumes, we observe that there are short periods of increased engagement. For Provider 3, this occurs between July 2023 and October 2023. Similarly, Provider 4 observes the highest volume of monthly tweets in July 2024. For Provider 2, monthly tweets consistently increase from May 2020 up until December 2020 before gradually starting to fall. However, Provider 1 observes significant variation in the number of monthly tweets tagging the PSP, lying anywhere between 500 and 1500 tweets in any given month of the study period. This pattern is interesting for several reasons. First, episodic spikes suggest that user attention toward PSPs on Twitter is event-driven rather than continuous. These surges are likely triggered by platform-specific shocks such as major outages, policy or user interface changes, large-scale promotional campaigns, regulatory announcements, or widely publicised fraud incidents. Second, the temporal asymmetry across platforms, specifically in the case of Provider 2, which also coincides with the initial phase of rapid adoption during COVID-19, suggests that social media engagement may be particularly intense during phases of onboarding, when users are still unfamiliar with the UPI system and more likely to seek help, report problems, or voice concerns publicly. As users adapt and platforms stabilise, engagement on Twitter starts declining. Ultimately, the patterns highlight that tweet volumes are not simply proxies for platform size or usage, but instead reflect moments of friction, transition, and contestation between users and PSPs. This makes monthly Twitter engagement a useful indicator of stress points in the digital payments ecosystem.

Figure 3 shows the volume of the top 3 classified issues each month from January 2024 to June 2025. This subset of the entire study period was chosen to aid visual clarity. Appendix 1.2 lists the top 3 classified issues for each month throughout the entire study period starting in January 2020.

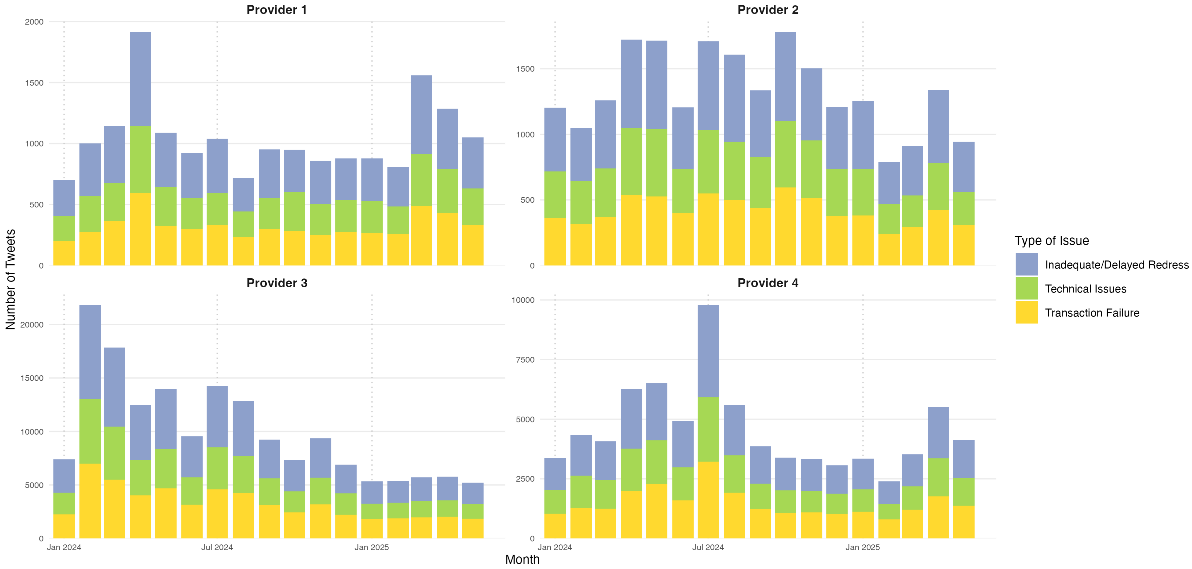


Figure 3: Top 3 Twitter Issues by Volume per Month

The top three issues that have been consistently dominating Twitter data are inadequate or delayed redress, transaction failure, and technical problems. Inadequate or delayed redress refers to situations where customers cannot access grievance redress channels, the redress mechanism fails to resolve their concerns effectively, or they do not receive redress in a timely manner. Transaction failure is defined as cases where money does not reach the intended recipient of a UPI transaction, including instances where funds are debited from the sender's bank account but not received by the recipient. Technical issues include difficulties in using the PSP's app due to technological problems like server failures or glitches that hinder the smooth completion of UPI transactions.

Given the nature of these issues, it is possible for the three to be interconnected. For example, when a transaction fails because of a technical problem with the PSP's app, customers may seek resolution through the PSP's redress mechanism. To explore these potential links, we calculate the pairwise correlations between the issues observed on Twitter. These are presented in Figure 4. This helps us confirm our hypothesis that there is some degree of correlation among the three types of customer issues.

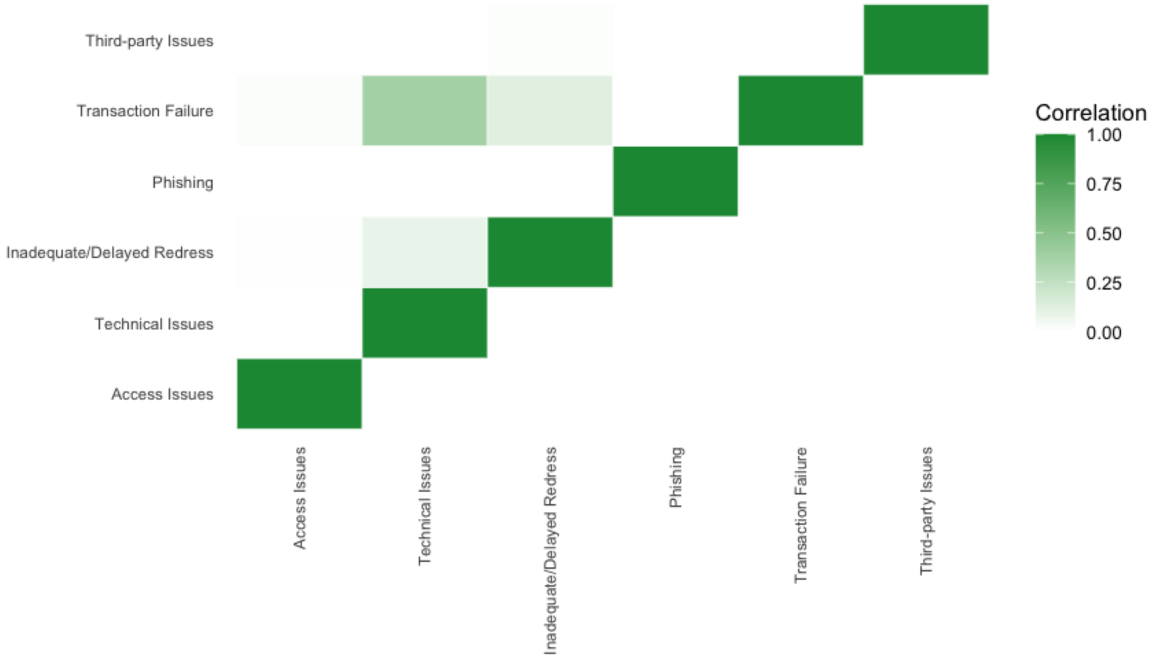


Figure 4: Pairwise Correlation between Twitter Issues

Additionally, we visualise the co-occurrence of different issues using a network graph in Figure 5. This is based on comprehensive data from January 2020 to June 2025. Network plots showing co-occurring issues for each year of the study are included in Appendix 1.3.

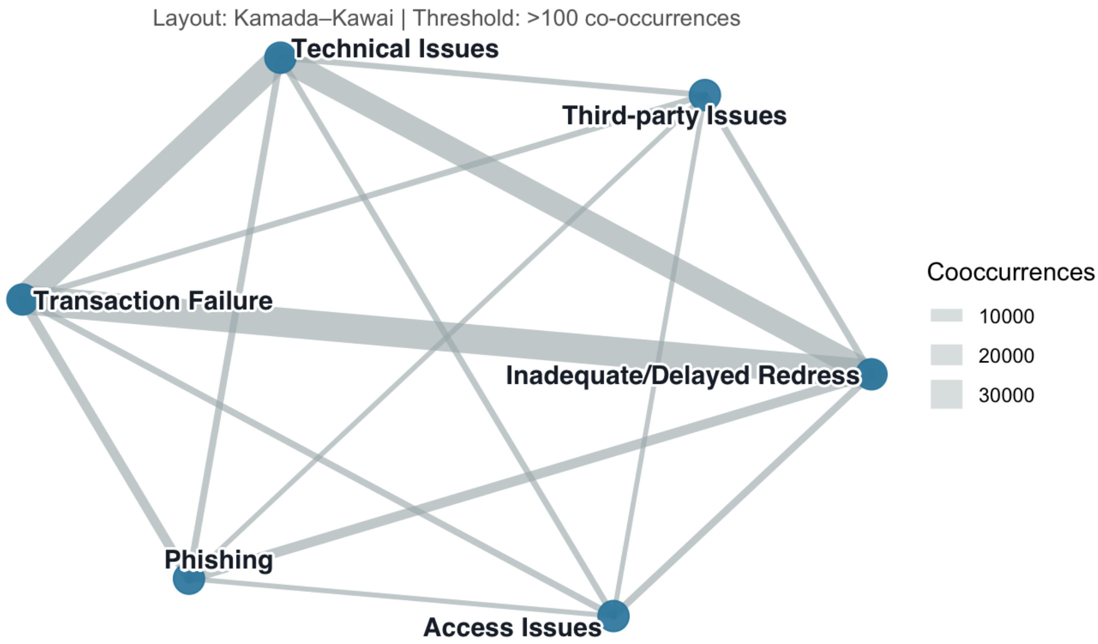


Figure 5: Network of Co-occurring Twitter Issues

In Figure 5, each issue is represented as a single node, and the lines connecting two nodes show that these issues occurred together in the same tweet. The width of each line indicates how often the issues co-occurred, based on the number of tweets where both issues appeared simultaneously. To highlight the strongest links, the network only includes issues with at least 100 co-occurrences. This threshold eliminates weaker or incidental connections and emphasises the most meaningful patterns of co-occurring issues in the Twitter data. From the figure, it is evident that there is some level of co-occurrence among all types of issues in the typology. This association is strongest between inadequate/delayed redress, transaction failure, and technical issues.

Lastly, Figure 6 shows an alluvial diagram which plots the share of all customer-facing issues in each quarter from Q1 2024 until Q2 2025. We select this subset from the entire study period to allow a cleaner visualisation of the diagram.

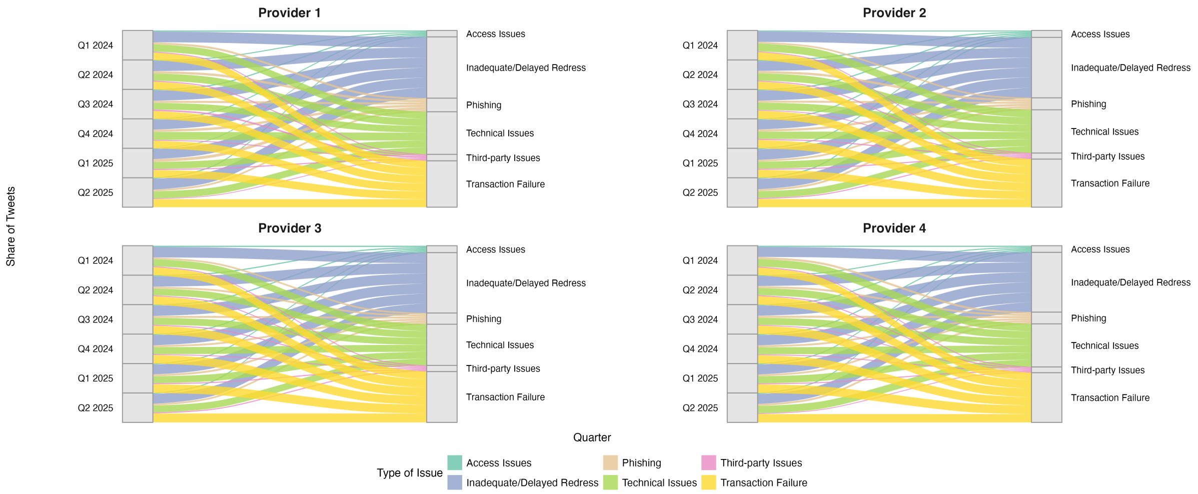


Figure 6: Quarter Flow of Twitter Issues

Figure 6 offers a visual summary of the relative frequency of different issue types reported on Twitter. Each issue is shown by a vertical block on the right-hand side of each plot in the panel, and the width of each stream indicates how often that issue appears in the data. The flows illustrate the distribution of issues across categories, highlighting which types of issues dominate the dataset and how their volume fluctuates across quarters (denoted on the left-hand side of each plot in the panel). Once again, inadequate or delayed redress, transaction failures, and technical issues stand out as the most pressing concerns on Twitter.

Part 2: Play Store

In this part, we discuss the findings from Play Store data. Figure 7 shows the monthly volume of Play Store reviews from January 2020 to June 2025. It should be noted that Play Store reviews for Provider 4 prior to September 2021 were not available on the Store and therefore were not collected. The complete daily volume data from January 2020 to June 2025 is included in Appendix 2.1. In this case, Provider 2 reported the highest number of daily reviews with an average of 107.17 reviews, followed by Provider 3, Provider 4 and Provider 1 at 75.93, 63.05 and 22.28 daily reviews respectively.

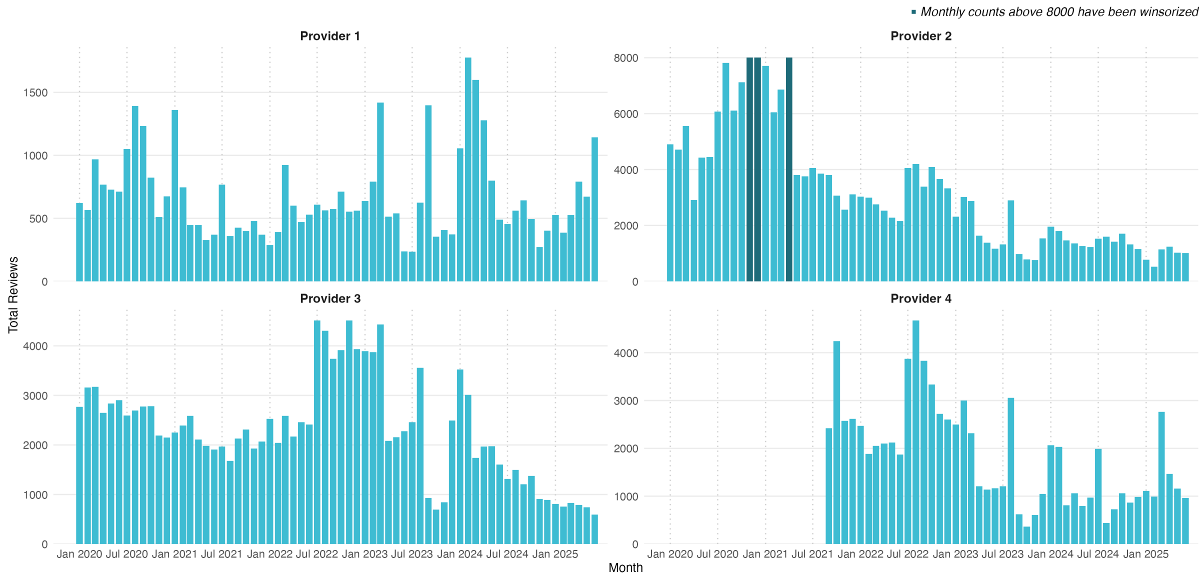


Figure 7: Monthly Volume of Play Store

Similar to Twitter, there is clear evidence of episodic spikes, temporal asymmetry, and volatility in the monthly volume of Play Store reviews. For Provider 2, we observe a significant decline in the monthly volume of reviews since May 2020. Meanwhile, Provider 3 and Provider 4 experience different periods across the study duration where the user's engagement fluctuates. Additionally, as observed for tweets referring to the Provider 1 handle on Twitter, Play Store registers a similar degree of volatility in terms of its monthly volume of reviews.

Figure 8 plots the volume of the top 3 classified issues each month from January 2024 to June 2025. Appendix 2.2 lists the top 3 classified issues for each month throughout the entire study period starting in January 2020.

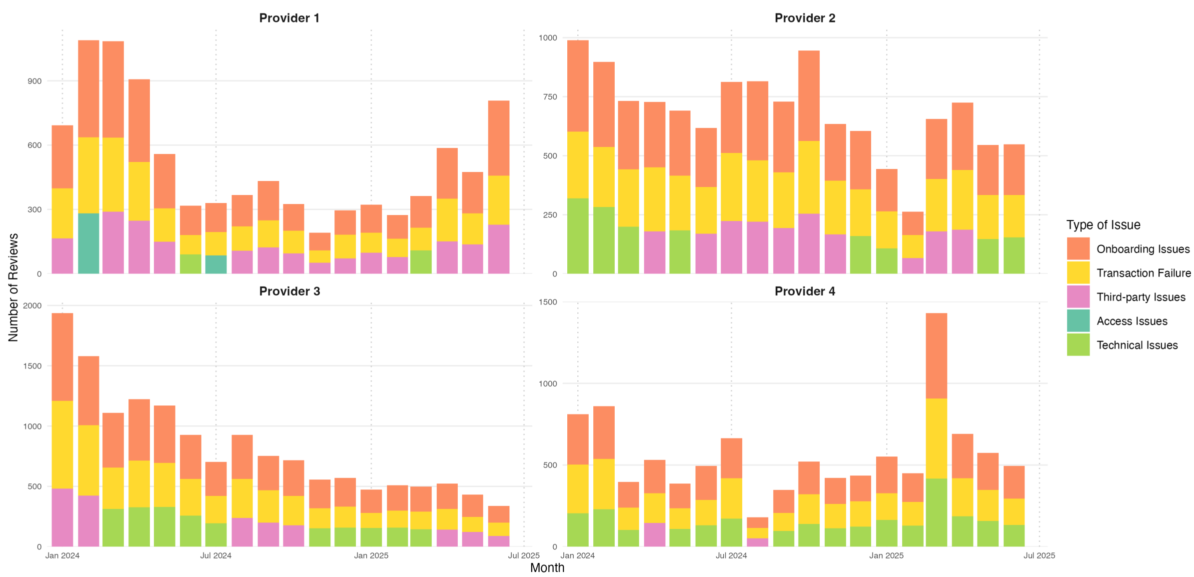


Figure 8: Top 3 Play Store Issues by Volume per Month

Both onboarding issues and transaction failure consistently stand out as the top two issues faced by customers between July 2024 and June 2025. Onboarding issues refer to the difficulties faced by users in signing up for the platform's service including hurdles faced while attempting to login to their UPI accounts. However, the third top issue varies by the platform. For example, Provider 3 sees third-party issues as a recurring problem across months while Provider 4 sees technical issues among top three on its platform. Third-party issues are defined as issues related to mobile wallets and transactions with third-party online merchants through UPI. For example, UPI can be used to recharge the user's FASTag accounts.

In Provider 3, there appear to be periods of alternation between third-party issues (January 2024 & February 2024; August 2024 to October 2024; April 2025 to June 2025) and technical issues (March 2024 to July 2024; November 2024 to March 2025). A similar trend can also be observed for Provider 2 where third-party issues occur from March 2024 to May 2024, July 2024 to November 2024, and then again February 2025 to April 2025. In the remaining months, this is replaced by technical issues, emerging from December 2024 to January 2025 and then again in May and June 2025.

This suggests there is some degree of correlation of customer issues across different platforms. Figure 9 shows the pairwise correlation of the different issues. Here, it can be observed that there is some degree of pairwise correlation between access issues, onboarding issues, third-party issues, inadequate/delayed redress as well as technical issues. This is an important insight, since it indicates that customers are using Play Store to register complaints on a wider range of issues, as compared to Twitter (see Figure 4 for reference).

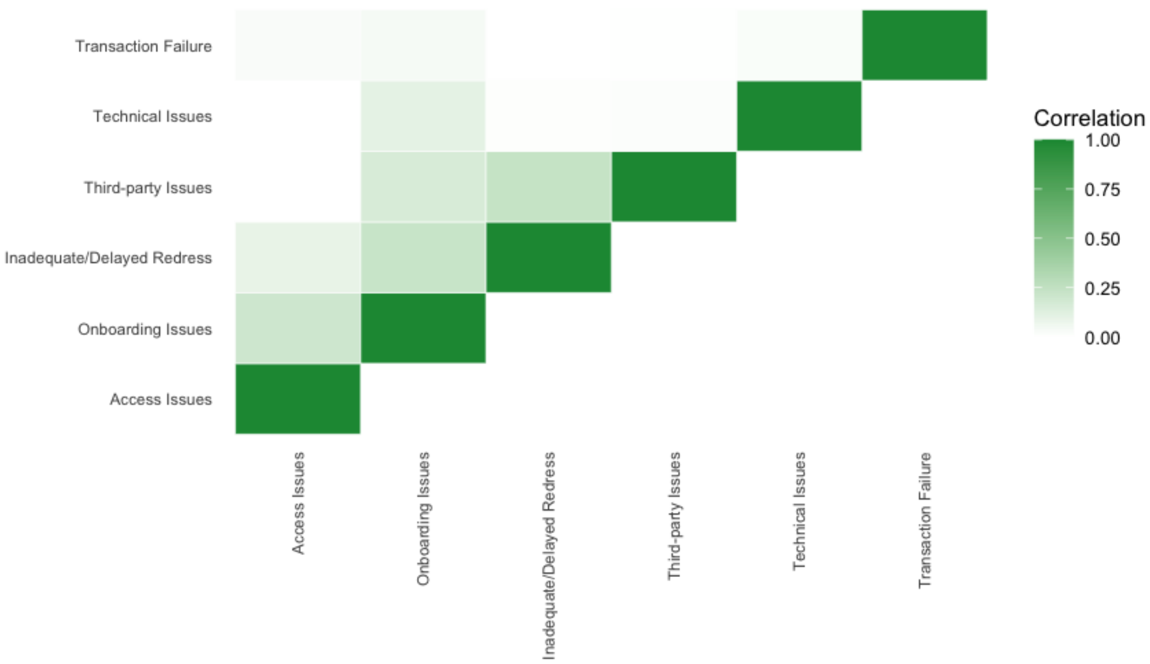


Figure 9: Pairwise Correlation between Play Store Issues

To investigate this further, Figure 10 plots the co-occurrence of different issues using a network graph. This is based on data from January 2020 to June 2025. Network plots showing co-occurring issues for each year of the study are included in Appendix 2.3.

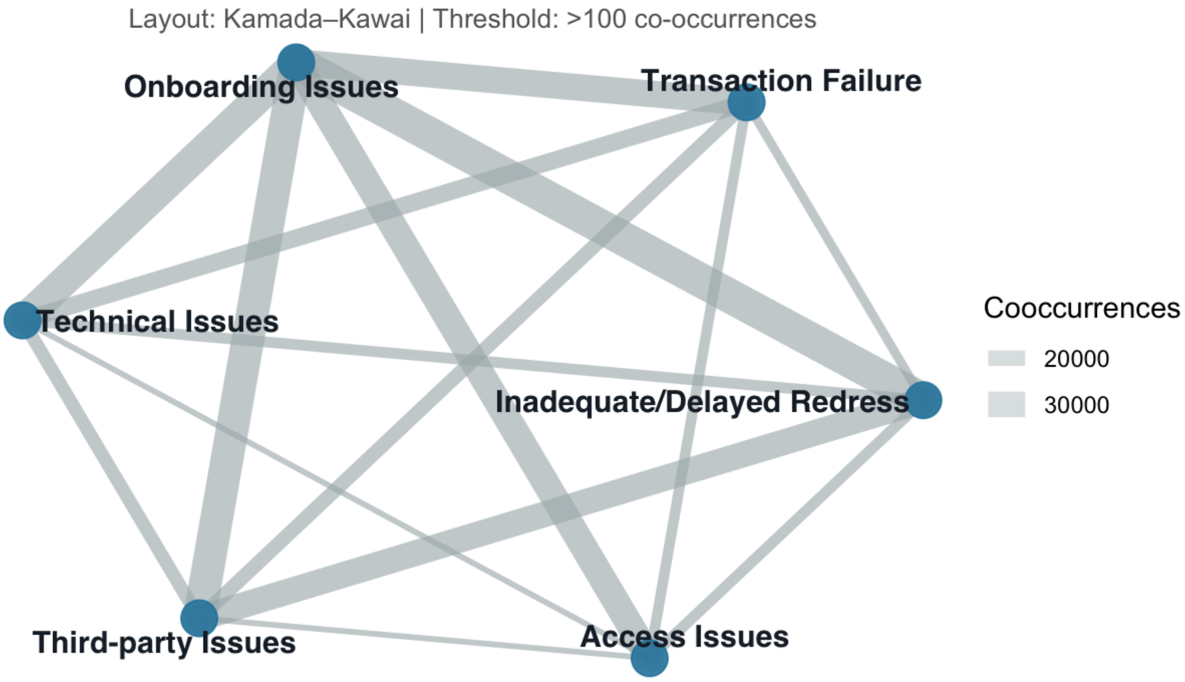


Figure 10: Network of Co-occurring Play Store Issues

It can be inferred from Figure 10 that customers of UPI are using Play Store to report onboarding issues since it displays the highest association with other kinds of issues that typically follow hurdles in accessing the UPI application. In other words, once this primary hurdle is resolved, customers are using Play Store reviews to report other associated issues such as technical issues, transaction failure, third-party issues, inadequate/delayed redress and access issues.

Lastly, Figure 11 shows an alluvial diagram which plots the share of all customer-facing issues in each quarter from Q1 2024 until Q2 2025.

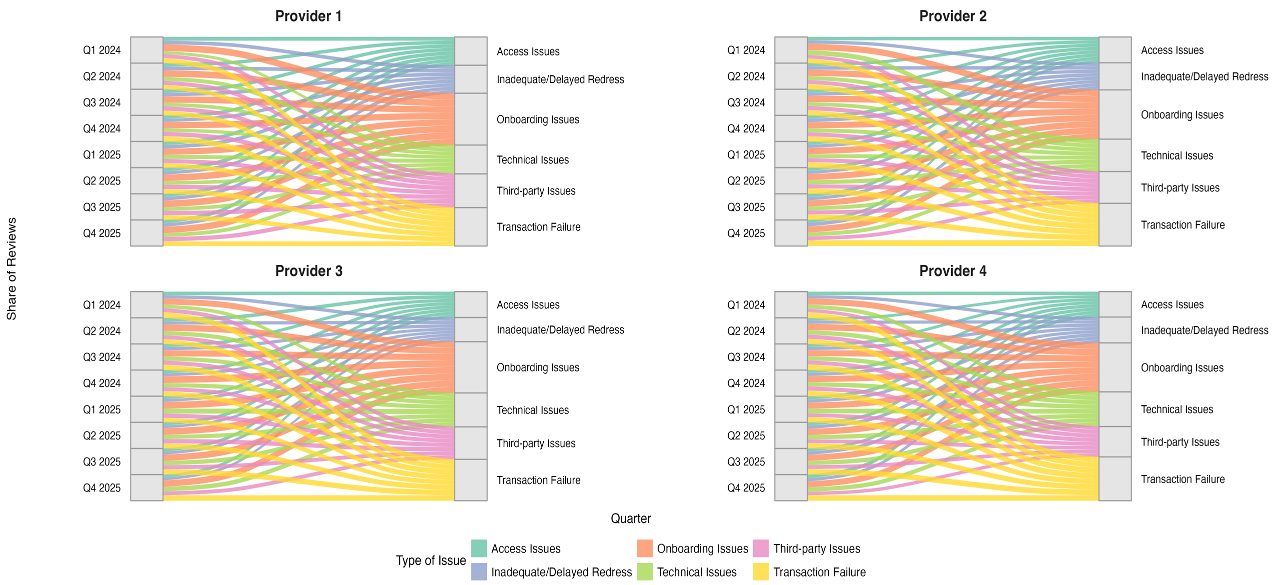


Figure 11: Quarter Flow of Play Store Issues

Similar to Figure 6, Figure 11 shows a summary of the relative frequency of different issue types reported on Play Store. This confirms our earlier hypothesis that customers are using Play Store to report a wider variety of issues faced by them while using the UPI application.

Concluding Remarks

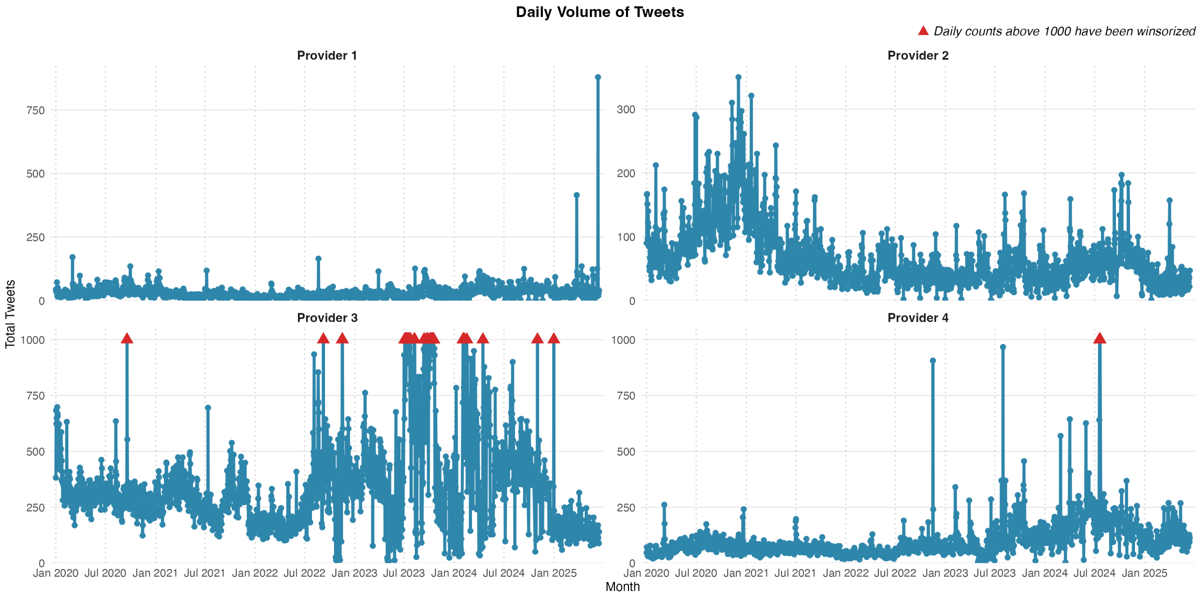
The objective of the report, as discussed earlier, was to demonstrate the types of information that market monitoring may uncover specifically in the context of financial customer protection in a high-frequency digital payments ecosystem like UPI.

By systematically tracking and classifying customer grievances expressed on social media, the study revealed that patterns of friction and unmet expectations that are not fully captured by formal grievance redress mechanisms can indeed be identified. The persistence and co-occurrence of issues such as transaction failure, technical problems, and inadequate or delayed redress across platforms point to structural challenges. At the same time, differences between Twitter and Play Store highlight how customers strategically choose platforms to voice distinct types of concerns, underscoring the need for multi-channel monitoring.

Taken together, the findings suggest that near-real-time market monitoring can equip PSPs and regulators with early warning signals, enable more responsive supervision, and ultimately strengthen trust and resilience in India’s digital payments infrastructure. Conversely, it enables PSPs to identify pain points that are not necessarily challenges for customer protection, but can be strengthened to enhance quality of service and customer satisfaction.

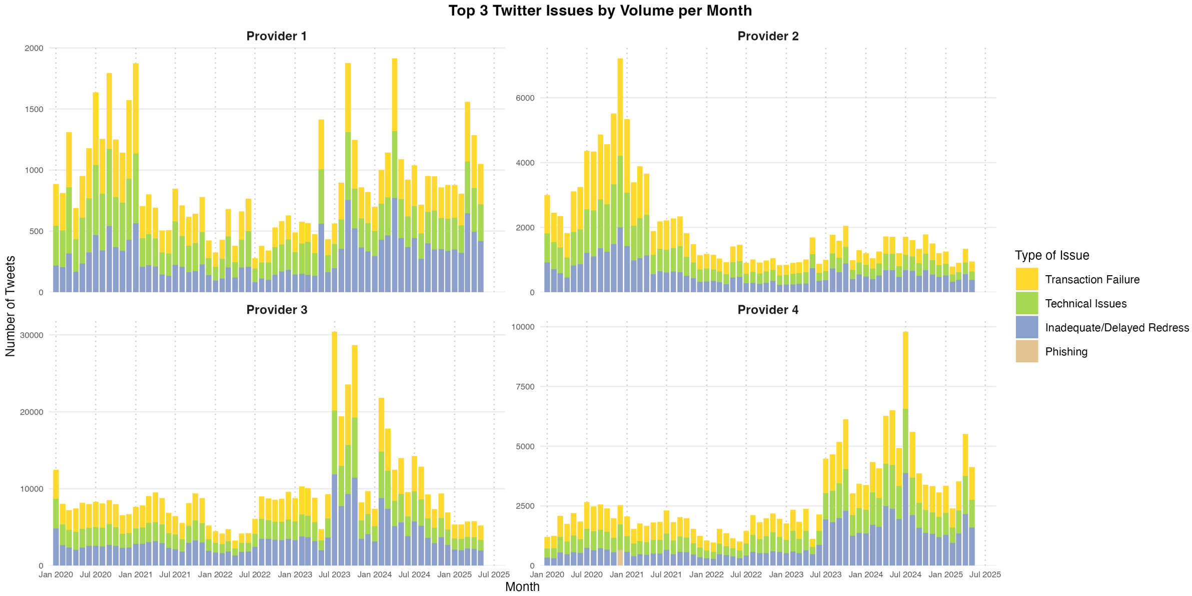
APPENDICES

Appendix 1.1: Daily Tweet Volume

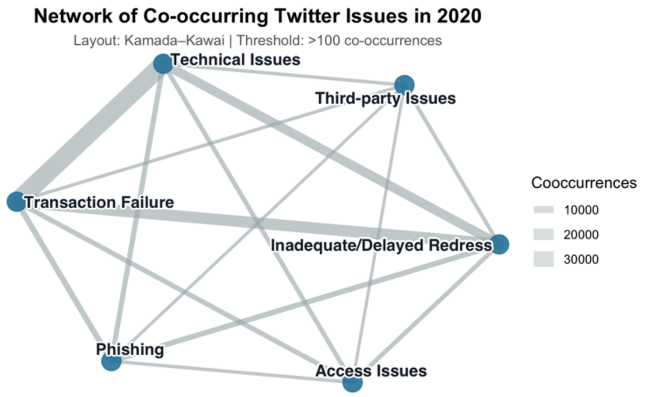


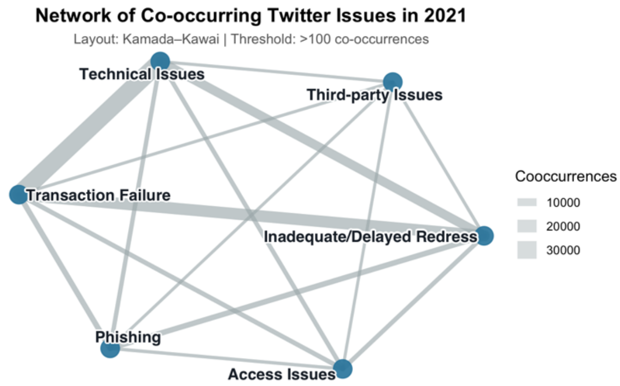
Daily counts above 500 have been winsorized to improve visualisation.

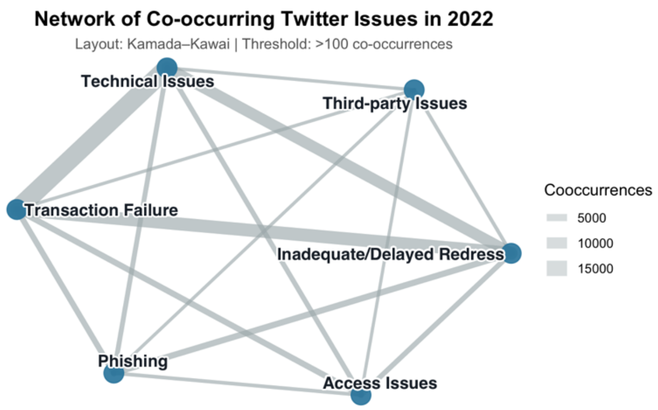
Appendix 1.2: Monthly Occurrence of user Issues on Twitter

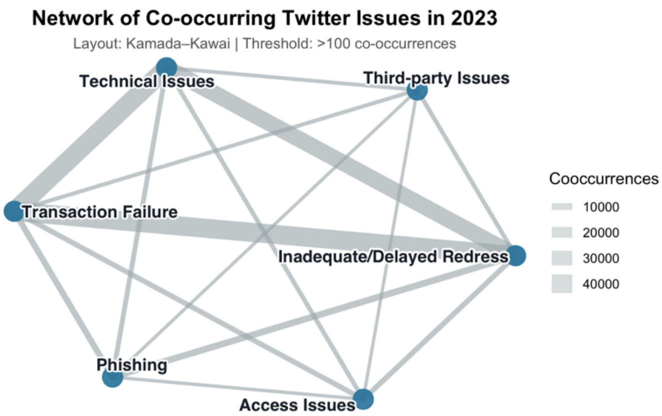


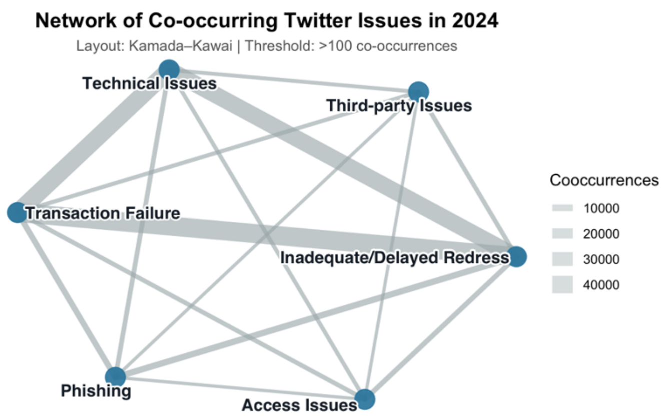
Appendix 1.3: Monthly Co-occurrence of user Issues on Twitter

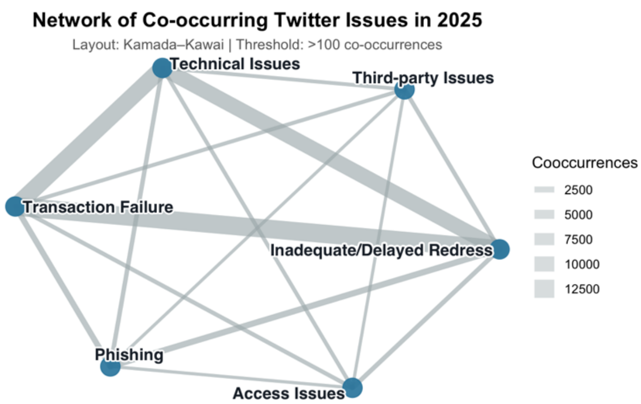




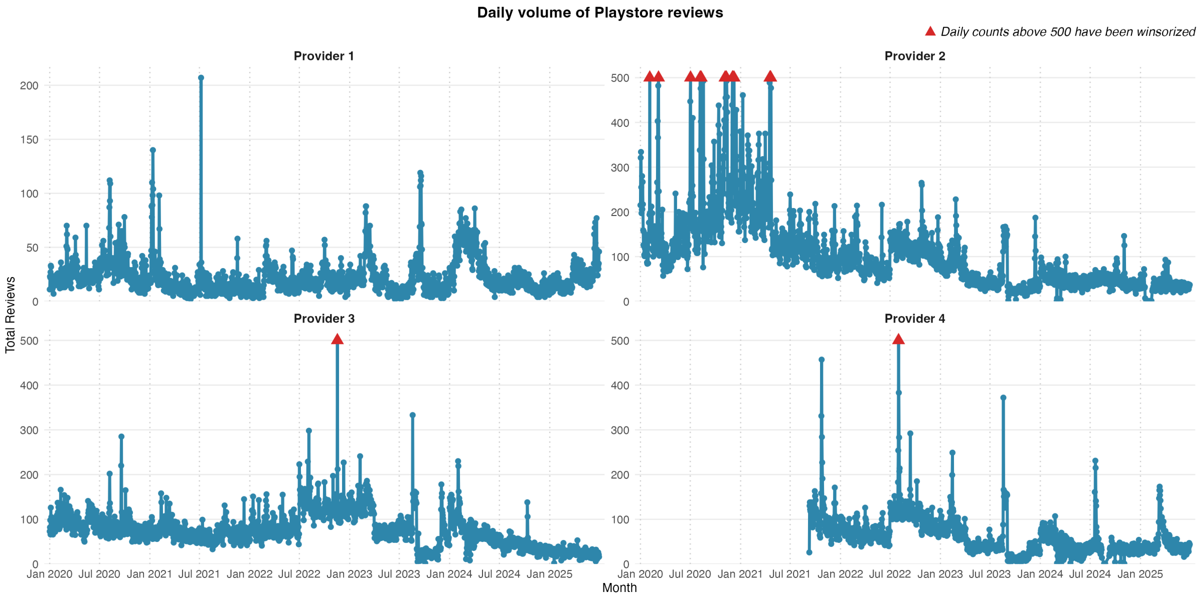






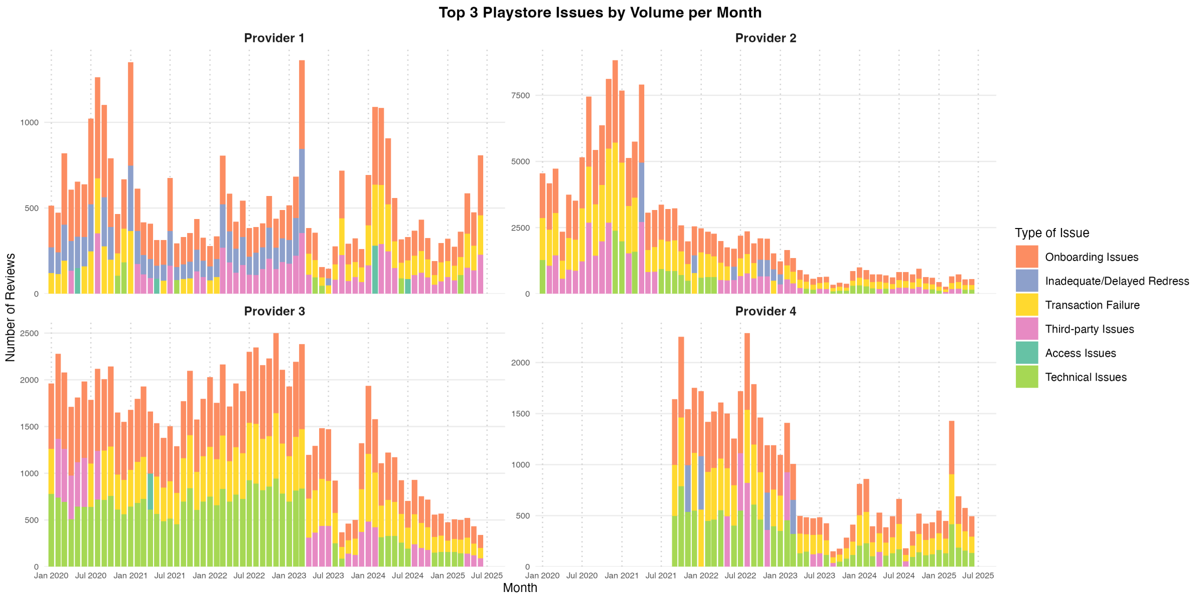


Appendix 2.1: Daily Review Volume



Daily counts above 500 have been winsorized to improve visualisation.

Appendix 2.2: Monthly Occurrence of user Issues on Play Store



Appendix 2.3: Monthly Co-occurrence of user Issues on Play Store