Fraud, Misrepresentation, and Transaction Security in Peer-to-Peer Trading Platforms: Detection Models and Policy Implications

Rizky Adinata[1] **and Bagas Wiratmoko**[2]

[1]**Universitas Teknologi Samudra Raya, Department of Computer Science and Engineering, Jl. Cendana No. 57, Mamuju, Sulawesi Barat, Indonesia**
[2]**Institut Informatika Nusantara Andalas, Department of Computer Systems and Networks, Jl. Sudirman No. 88, Payakumbuh, Sumatera Barat, Indonesia**

Sciencespress is a publisher of peer-reviewed scientific journals, established in 2018 with a mission to advance global research dissemination. Specializing in multidisciplinary fields such as life sciences, environmental research, and technology, the platform emphasizes rigorous peer review to maintain high academic standards.

RESEARCH ARTICLE

Abstract

Peer-to-peer trading platforms have expanded the set of transactions that occur without traditional intermediaries, relying instead on software-mediated trust, lightweight identity signals, and platform governance. As participation grows and cross-border trading becomes routine, the same features that improve market access also increase exposure to fraud, strategic misrepresentation, and disputes with limited offline enforceability. This paper analyzes detection and security mechanisms for peer-to-peer trading under adaptive adversaries, focusing on how data-driven models interact with product design, transaction protocols, and policy constraints. A technical framework is developed that treats fraud risk as a sequential, partially observed process spanning onboarding, listing, negotiation, payment, fulfillment, and dispute resolution, with feedback loops created by enforcement actions and reputation systems. The paper examines learning objectives aligned to operational costs, including chargebacks, subsidy leakage, account recovery, manual review capacity, and user attrition, and it discusses how to manage label noise, delayed outcomes, and selection bias induced by interventions. Modeling approaches include cost-sensitive classification, semi-supervised anomaly detection, graph-based inference over user-transaction-device networks, and decision-focused thresholding with calibration and uncertainty. Transaction security is discussed as a complement to detection, emphasizing escrow-like holds, authenticated messaging, evidence capture, and settlement design that reshapes incentives. Policy implications are derived for transparency, appeals, privacy, cross-jurisdiction enforcement, and proportional sanctions, highlighting conditions under which model governance and protocol choices reduce harm without suppressing legitimate trade.

1 Introduction

Peer-to-peer trading platforms mediate exchanges between individuals and small merchants across a wide range of goods and services, including secondhand marketplaces, local pickup listings, digital goods, and peer-facilitated payments [1]. These platforms typically operate with limited ex ante verification relative to traditional retail or banking rails, substituting user ratings, lightweight identity checks, and platform rules for the institutional controls that historically constrained counterparty risk. This substitution is economically attractive because it reduces friction and increases market thickness, yet it also creates an environment where opportunistic behavior can be profitable at scale. Fraud in this context is not only direct theft through payment reversal or non-delivery, but also strategic misrepresentation of identity, product condition, provenance, or intent, as well as exploitation of platform subsidies, dispute processes, and promotional mechanisms.

OPEN ACCESS
Reproducible Model

Edited by
Associate Editor

Curated by
The Editor-in-Chief

A defining challenge for peer-to-peer settings is that adversarial behavior is endogenous to platform design. When a platform changes its verification, escrow, ranking, or enforcement policy,

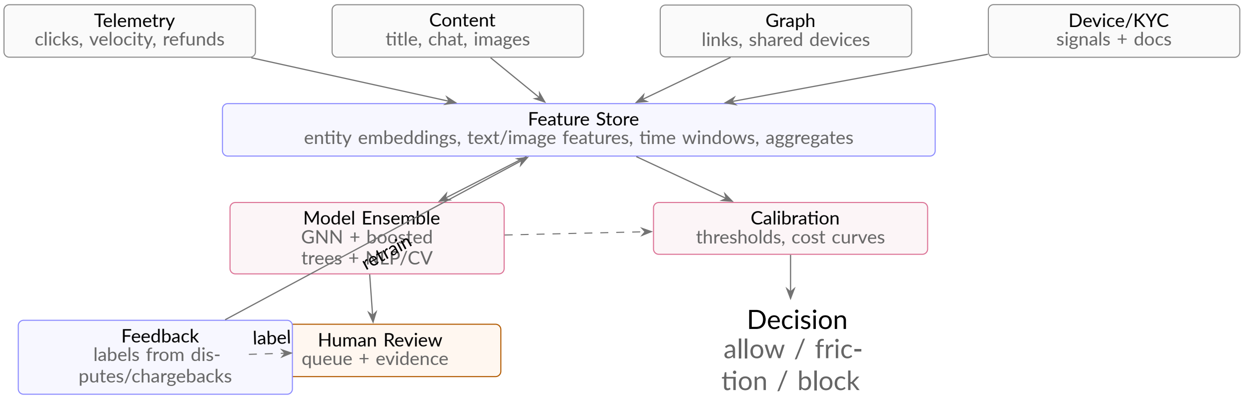

Figure 1. End-to-end fraud detection pipeline. Heterogeneous signals (telemetry, content, graph structure, and device/KYC) are merged into a feature store feeding an ensemble of specialized models. Calibrated outputs drive risk actions (allow, friction, or block), while reviewer decisions and downstream outcomes (disputes, chargebacks) close the loop via labeling and retraining.

Type of fraud	Description	Key signal	Typical severity
Payment fraud	Chargebacks, fake payment proof	Inconsistent payment timestamps	High (direct monetary loss)
Identity spoofing	Stolen / synthetic identities	Multiple accounts sharing KYC fields	High (platform trust erosion)
Product misrepresentation	Item not as described	High dispute rate per seller	Medium (reputational + refund costs)
Service non-delivery	No delivery after payment	Long unresolved order age	High (user churn)
Collusive feedback	Fake reviews, rating inflation	Dense rating cliques	Medium (ranking distortion)
Account takeover	Compromised legitimate accounts	Sudden device / IP change	Critical (hard to detect early)

Table 1. Main fraud patterns observed on peer-to-peer trading platforms.

adversaries adapt their tactics, often shifting from easily detectable single-account abuse to coordinated behavior across multiple accounts, devices, and payment instruments. As a result, platform risk management must be understood as a dynamic control problem with partial observability, constrained by privacy obligations, legal requirements, user experience targets, and operational budgets for manual review. Detection models are necessary but not sufficient; they operate within a broader system that includes identity and device controls, payment and settlement protocols, messaging and content moderation, reputation aggregation, and dispute resolution [2]. Each subsystem produces signals that can improve inference, but each also creates new attack surfaces and incentives.

This paper develops a technical and policy-oriented account of fraud, misrepresentation, and transaction security in peer-to-peer trading. The emphasis is on detection models that are operationally meaningful, meaning that their outputs can be translated into actions such as step-up verification, payment holds, listing throttles, warnings, or account restrictions, and on how those actions interact with user behavior and market outcomes. The paper treats fraud risk as a lifecycle phenomenon, where early-stage signals during onboarding and listing creation are different in character from late-stage signals arising during fulfillment and disputes. It also treats misrepresentation as a continuum ranging from ambiguous quality descriptions to deliberate counterfeiting and identity deception, with implications for evidentiary standards and remedies.

A key practical difficulty is that ground truth labels for fraud are noisy and delayed. Many harmful events are never reported, and many reported events remain ambiguous because evidence is incomplete or contested [3]. Chargebacks, delivery failures, and user complaints are informative but imperfect proxies for intentional wrongdoing. Moreover, platform interventions alter what data is observed: if a model blocks suspicious transactions, the platform may never learn whether they would have been fraudulent, creating selection bias that can degrade future models if not explicitly addressed. This paper therefore focuses on modeling choices that are robust to label noise, censoring, and feedback, including calibrated risk scores, causal approaches to policy evaluation, and online learning under drift.

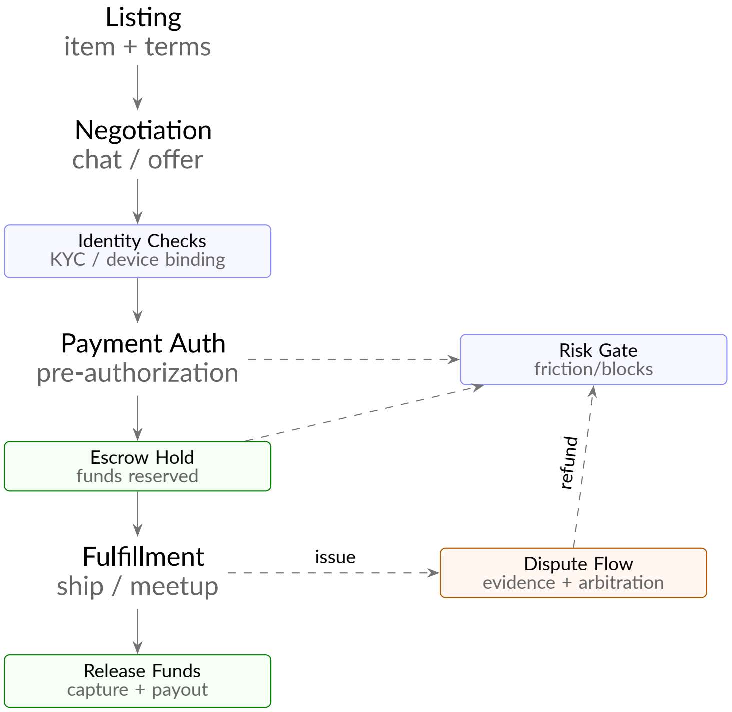

Figure 2. Secure transaction lifecycle emphasizing escrow and intervention points. Identity checks and risk gates can introduce friction before authorization or during escrow. Funds are released only after fulfillment confirmation, while dispute handling (evidence collection and arbitration) provides structured remediation for misrepresentation, delivery failures, and payment reversals.

Feature group	Example features	Detection rationale
Identity	KYC mismatch score, account age	Distinguish genuine from synthetic identities
Device	Device fingerprint entropy, OS changes	Surface shared or rapidly rotating devices
Network	IP ASN, geolocation hops, proxy score	Highlight suspicious access networks
Behavioral	Inter-arrival times, click paths	Capture bots and scripted trade flows
Content	Text embeddings, image hashes	Identify misrepresentation and phishing patterns
Platform	Dispute history, past sanctions	Encode institutional memory into risk scoring

Table 2. Feature families used in the fraud detection models.

The policy dimension is inseparable from the technical one. Platforms operate within consumer protection regimes, financial crime obligations, privacy laws, and, increasingly, governance expectations around automated decision-making. Decisions such as account suspension or fund withholding have due-process implications, and the distribution of false positives and false negatives can raise fairness concerns, particularly when risk signals correlate with protected attributes or with socioeconomic proxies. Additionally, cross-border peer-to-peer trading complicates enforcement because legal remedies and reporting channels vary by jurisdiction [4]. Policy constraints can therefore reshape the feasible set of model features, retention periods, and evidence-handling practices, which in turn influences detection performance and user trust.

The analysis proceeds by characterizing the threat landscape and the lifecycle of fraud and misrepresentation in peer-to-peer trading, then by detailing data sources and measurement problems that affect learning systems. It then develops detection and risk scoring models with explicit operational objectives and discusses transaction security architectures that change incentives and improve evidentiary quality. Finally, it draws policy implications for governance, transparency, and proportional enforcement, emphasizing how to align platform incentives with user protection while preserving legitimate commerce.

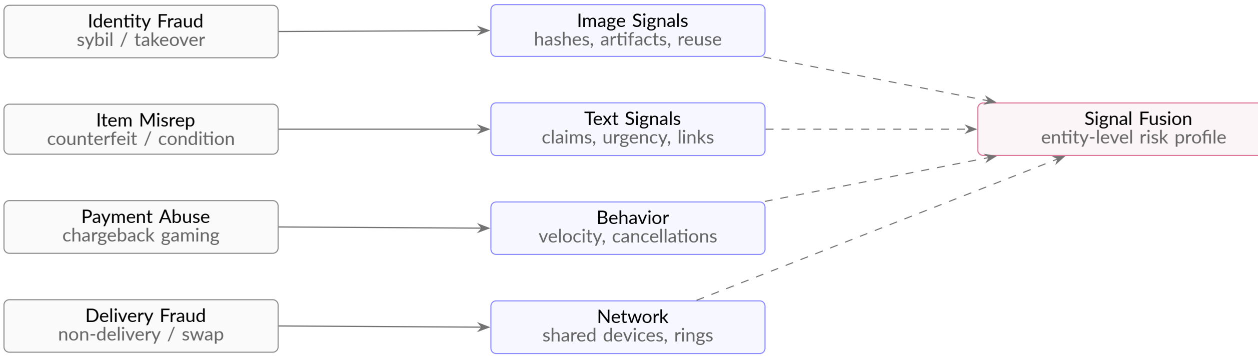

Figure 3. Misrepresentation and fraud taxonomy mapped to observable signals. Core abuse categories (identity, item misrepresentation, payment abuse, and delivery fraud) manifest through different evidence channels. Text/image features, behavioral telemetry, and network structure are fused into a unified entity-level view to support consistent scoring across accounts, listings, and transactions.

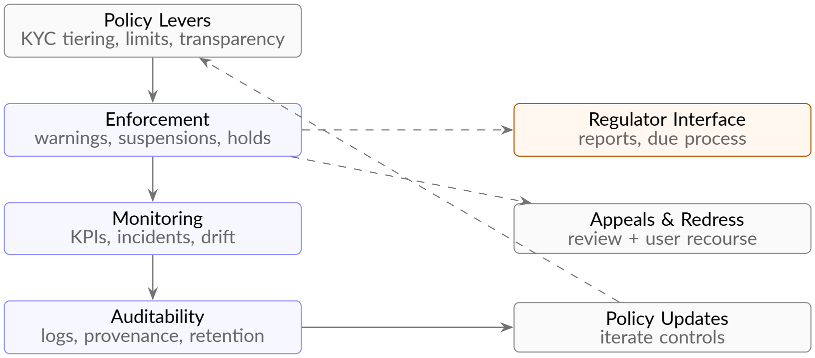

Figure 4. Governance loop linking detection outcomes to enforceable platform policy. Policy levers define guardrails and proportional frictions; enforcement actions are monitored through operational metrics and incident reviews; auditability supports accountability and evidence-based decision making. Appeals and regulator-facing reporting encourage due process, and feedback from audits drives iterative policy updates.

2 Fraud and Misrepresentation in the Peer-to-Peer Transaction Lifecycle

Fraud and misrepresentation on peer-to-peer platforms can be understood as strategic behavior that exploits information asymmetry, weak identity binding, limited enforcement, and the platform's need to balance friction against safety. While the specific manifestations vary by domain, a common structure is that the platform mediates discovery and communication, often provides payment rails or payment facilitation, and offers a dispute process that reallocates losses when parties disagree. Attackers aim to extract value by causing the platform or the counterparty to bear losses, and misrepresentation aims to increase the attacker's expected payoff by shifting beliefs about quality, intent, or identity [5].

The lifecycle begins at onboarding, where accounts are created and initial trust signals are established. Weak identity binding enables rapid re-entry after enforcement and facilitates coordinated behavior, while strong identity checks can deter legitimate participation and raise privacy concerns. The adversary's optimization often includes account acquisition cost, expected lifespan, and the value extractable per account. When platforms deploy step-up verification, attackers can respond by specializing: some accounts are used for reputation building, some for high-risk transactions, and some for laundering reputational signals through interactions that appear legitimate.

The listing or offer creation stage introduces content and metadata signals, including product descriptions, images, pricing, and category selection. Misrepresentation here includes overstating

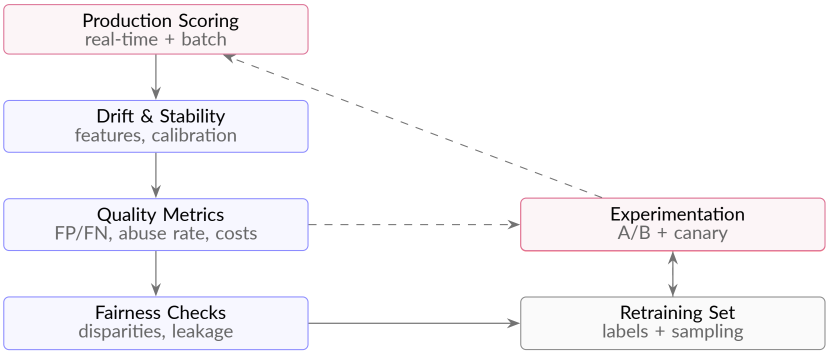

Figure 5. Operational monitoring and continuous learning for deployed fraud models. Production scoring is tracked for drift, calibration stability, and outcome quality (including cost-sensitive error tradeoffs). Fairness checks and incident response workflows feed curated retraining sets, while experimentation (A/B and canary releases) enables controlled updates to reduce abuse without degrading user experience.

Platform	Users (M)	Transactions (M)	Confirmed fraud rate (%)
Alpha (goods)	8.4	96.2	0.42
Beta (services)	3.1	37.8	0.67
Gamma (assets)	1.7	21.4	1.25
Combined	13.2	155.4	0.61

Table 3. Descriptive statistics of the peer-to-peer trading datasets.

condition, omitting defects, mislabeling authenticity, or presenting stolen goods as legitimate. Fraudulent listings may be designed to trigger high conversion through underpricing, scarcity cues, or high-demand items [6]. The technical challenge is that legitimate sellers can also exhibit unusual patterns, especially during seasonal demand shocks or when a seller is liquidating inventory. Pure anomaly detection can therefore be brittle without context.

Negotiation and messaging form a major channel for manipulation. Off-platform redirection, pressure tactics, and requests for alternative payment methods can indicate elevated risk, yet platforms must consider privacy and free-expression norms when analyzing messages. Even when message scanning is permissible, adversaries can obfuscate intent through coded language, images, or staged conversations. From a modeling perspective, the messaging phase is important because it produces behavioral sequences that reflect intent more than static profile features do, but it also raises the highest sensitivity in governance because it touches user communications.

Payment and settlement choices determine the economic feasibility of many fraud strategies [7]. If payment is reversible and platform protections are weak, the platform may be exposed to chargeback fraud, unauthorized payment instrument use, or disputes that exploit ambiguity. If payment is irreversible, the buyer's risk increases, which can reduce transaction volume unless escrow-like protections exist. In peer-to-peer contexts, a platform often must decide whether to act as merchant of record, whether to hold funds pending confirmation, and how to allocate liability across buyers, sellers, and the platform. These decisions shape the attacker's expected return and thus the prevalence of certain tactics.

Fulfillment and delivery add further uncertainty. For shipped goods, proof-of-delivery is imperfect, and address fraud can be used to create confusion. For local meetups, physical handoff reduces certain risks but introduces safety concerns and evidentiary gaps [8]. Digital goods and services can be delivered instantly, increasing the speed at which attackers can extract value and reducing the time window for detection. Disputes arise when one party claims non-delivery, misdescription, or unauthorized use. The platform's dispute policy becomes a target: if it reliably favors buyers, then sellers can be defrauded via false claims; if it reliably favors sellers, then

Model	ROC-AUC	F1 (fraud class)	PR-AUC (fraud)
Logistic regression	0.876	0.322	0.214
Random forest	0.911	0.387	0.263
XGBoost	0.933	0.425	0.294
Graph neural network	0.947	0.451	0.318
Transformer (sequence)	0.952	0.463	0.331
Hybrid (GNN + Transformer)	**0.963**	**0.489**	**0.356**

Table 4. Comparison of supervised fraud detection models on the combined dataset.

Removed feature block	ΔROC-AUC	ΔPR-AUC	Interpretation
Content features	-0.011	-0.018	Text and images are crucial for misrepresentation cases
Behavioral features	-0.019	-0.027	Temporal activity patterns carry strong predictive signal
Network features	-0.024	-0.031	Shared IP and graph signals matter for collusion
Device features	-0.009	-0.012	Helps uncover multi-accounting at scale
User reports / disputes	-0.006	-0.008	Human feedback refines borderline decisions

Table 5. Ablation study of feature groups in the hybrid detection model.

buyers can be defrauded via non-delivery or counterfeit delivery.

Reputation and feedback systems are intended to mitigate information asymmetry but can be gamed. Self-dealing, reciprocal rating rings, and strategic timing can inflate apparent trust. A key difficulty is that reputation is a function of observed outcomes, which are themselves influenced by who is willing to transact with whom. If a platform's ranking system boosts accounts with high conversion and low complaint rates, attackers can build a benign history through low-risk transactions before switching to high-value fraud, a pattern that creates temporal non-stationarity that detection models must anticipate [9].

Misrepresentation differs from fraud in that it often exists on a spectrum of interpretability, and evidence may be subjective. A buyer may interpret a product as defective while a seller claims it is within expectations. This ambiguity can be exploited by adversaries who operate near the boundary of enforceable rules, using plausible deniability to reduce sanction risk. Platforms therefore need decision frameworks that incorporate uncertainty and that select remedies proportional to confidence, such as partial refunds, warnings, or additional verification, rather than binary outcomes only.

Coordination and collusion represent higher-order threats. When platforms offer promotions, referral bonuses, fee waivers, or shipping subsidies, coordinated groups can extract value through cyclical transactions that appear legitimate on the surface. Similarly, laundering of reputational signals can be done through networks of accounts that trade low-value items to generate a history of completed transactions [10]. These behaviors are difficult to detect using only per-account aggregates, motivating graph-based modeling across users, devices, payment instruments, addresses, and transaction partners.

A final element of the threat landscape is adversarial adaptation to detection. When features or rules become predictable, attackers can learn to mimic benign patterns. This creates an ongoing arms race in which the platform must invest in feature hardening, model monitoring, and security design that reduces the information attackers can glean about enforcement. At the same time, platforms must provide enough transparency to maintain user trust and comply with governance expectations. The resulting tension motivates designs that separate internal high-dimensional risk signals from user-facing explanations that are accurate but not easily weaponized.

3 Data, Measurement, and Labeling Under Feedback and Uncertainty

Detection and policy evaluation depend on data, yet peer-to-peer trading data is heterogeneous, noisy, and shaped by platform interventions [11]. A platform typically observes account at-

Operating point	TPR (%)	FPR (%)	Expected loss / 10k trades (USD)
Lenient (low threshold)	94.1	4.8	3,420
Balanced	88.7	2.1	2,160
Strict (high threshold)	77.5	0.9	1,540
High-friction only	70.3	0.5	1,480
Manual-review heavy	90.2	1.2	1,310

Table 6. Trade-off between operating points, false alarms, and expected platform loss.

Dimension	Real-time scoring	Batch scoring	Remarks
Latency	< 150 ms	Minutes to hours	Real-time protects in-flight trades
Compute cost	High per event	Lower amortized cost	Batch better for large-scale retrospection
Coverage	Subset of features	Full historical feature set	Batch uses richer context
Model complexity	Moderately constrained	Less constrained	Heavy models scheduled off-peak
Typical use cases	Pre-trade risk, login checks	Retroactive sweeps, model retraining	Complementary deployment modes

Table 7. Operational comparison of real-time and batch fraud detection pipelines.

tributes, device and network metadata, behavioral logs, listing content, message interactions, payment events, fulfillment status, and dispute outcomes. Each modality has distinct reliability and sensitivity. Identity signals can be stable but privacy-sensitive, device fingerprints can be high value but may drift or be regulated, and complaint text can be rich but linguistically ambiguous. The challenge is to construct features that are predictive, robust, and compliant with constraints on data usage and retention.

The core supervised learning obstacle is label quality. A common positive label is derived from confirmed fraud outcomes, such as chargebacks attributed to fraud, policy violations adjudicated as scams, or repeated complaints with consistent evidence. However, many fraudulent transactions never produce such labels because victims do not report, because they resolve privately, or because the loss is too small [12]. Negative labels are even more problematic because an absence of complaints does not imply absence of misconduct. This creates asymmetric label noise and a positive-unlabeled learning setting where labeled positives are a biased sample of the true positives. The bias is correlated with user sophistication, transaction value, and the platform's own dispute policy, which affects reporting incentives.

Labels are also delayed. Chargebacks and disputes can arise days or weeks after a transaction, and some misrepresentation is discovered only after extended use. Delayed labels imply that training data reflects the past state of the platform and the past attacker strategy distribution. Drift is therefore structural rather than incidental [13]. Additionally, platform actions such as blocking a transaction or requiring additional verification change which outcomes are observed. If a high-risk transaction is blocked, the platform does not observe whether it would have resulted in fraud, and if additional verification is required, the subset of users who proceed may differ systematically from those who abandon, producing post-intervention selection bias.

Measurement also depends on operational definitions. Fraud can mean unauthorized payment, non-delivery with intent, counterfeit goods, or abusive use of refunds. Misrepresentation can include borderline cases. Different teams may label the same case differently depending on policy goals, which can introduce label inconsistency. For technical models, it is often useful to define multiple target variables aligned to action types, such as probability of chargeback, probability of dispute escalation, probability of counterfeit claim, or probability of user harm as measured by complaint severity [14]. Multi-task modeling can then share representations while allowing decision-specific calibration.

Feature construction must account for adversarial manipulation. Profile features like username patterns, listing price anomalies, or transaction velocity can be predictive but may be easily mimicked. Behavioral sequences such as time-to-first-message, ratio of initiated to completed transactions, or escalation patterns can be harder to forge consistently, yet they can still be manipulated by coordinated groups. Network features across shared devices, addresses, payment

Policy area	Instrument	Target actor	Expected effect
Identity assurance	Tiered KYC with risk-based triggers	High-volume traders	Reduces synthetic and mule accounts
Dispute handling	Time-bounded resolution protocols	Platforms and mediators	Lowers unresolved claims and churn
Transparency	Standardized risk notices	Buyers and sellers	Aligns expectations on platform protections
Sanctions	Graduated penalties and bans	Repeat offenders	Deterrence via predictable consequences
Education	In-app security tutorials	New users	Cuts down on avoidable victimization
Data sharing	Industry threat-intel exchanges	Platforms and regulators	Faster cross-platform takedown of rings

Table 8. Regulatory and platform policy tools relevant to peer-to-peer fraud.

Nudge type	Click-through rate (%)	Drop in disputed trades (%)	Notes
Friction banner at checkout	41.8	6.2	Short warning for high-risk counterparties
Pre-trade checklist	35.4	8.9	Encourages use of escrow and in-app chat
Counterparty risk score display	57.1	11.3	Simple visual score near username
Delayed payout reminder	62.6	9.7	Highlights benefits of holding funds in escrow
Post-trade survey prompt	24.3	3.4	Provides extra labels for future models

Table 9. Experimental results from user-facing security nudges on a pilot platform.

tokens, or counterparties are valuable because they capture structural constraints on attacker operations, but they risk false associations in shared environments such as households, campuses, or public networks.

Text and image data from listings and messages can provide high-resolution signals for misrepresentation, such as template reuse, inconsistent metadata, or semantically suspicious claims. However, content models must be robust to benign linguistic variation and to evolving slang [15]. Furthermore, content analysis can increase privacy sensitivity, especially for private messages. A governance approach is often to use content-derived features that are narrowly scoped to policy enforcement, to store derived embeddings rather than raw content when permissible, and to implement strict access controls and retention limits. Even when technically feasible, these choices influence model performance and auditability.

Ground truth for misrepresentation often requires external verification. Authenticity of goods, for example, may not be reliably determined from platform data alone. Platforms can use third-party verification services or require serial numbers or provenance documents, but this increases friction and can exclude legitimate sellers. An alternative is to treat authenticity as a probabilistic variable informed by category risk, seller history, and complaint patterns, and to prioritize interventions that reduce harm, such as escrow holds for high-risk categories, rather than attempting definitive classification in all cases [16].

Dispute resolution processes generate rich labels but also embed policy choices. If a platform's policy strongly favors one side, the observed dispute outcomes reflect that bias. A model trained to predict dispute outcomes can inadvertently learn to reproduce policy bias rather than true misconduct. Separating the prediction of objective events, such as delivery confirmation or payment reversal likelihood, from subjective adjudication outcomes can help, but often the platform must operate with imperfect proxies. A practical approach is to maintain distinct datasets for operational risk prediction and for policy compliance prediction, and to periodically recalibrate models when policies change.

Another measurement issue is the definition of harm. Fraud detection is often framed as minimizing financial loss, but peer-to-peer trading involves non-financial harms such as time loss, emotional distress, and safety concerns for in-person meetups [17]. These harms are difficult to quantify. Platforms may use complaint severity taxonomies or user surveys to estimate non-financial harm, but these are again biased by reporting. A model that optimizes purely for financial loss may under-protect low-value transactions where the user experience cost is nonetheless significant. Incorporating user-reported dissatisfaction and churn signals can partially address this, though these signals are confounded by many non-fraud factors.

Operational constraints shape labeling as well. Manual review capacity is limited, so only a subset of cases are investigated. This creates a feedback loop where the model's prior decisions determine which cases are reviewed and thus which labels are generated [18]. Active learning can help by selecting uncertain or informative cases for review, but it must be designed carefully to avoid over-sampling certain user groups or transaction types. Additionally, evidence collection must be standardized; otherwise labels will reflect reviewer effort rather than underlying reality.

Finally, privacy and compliance constraints can limit data availability and retention, increasing the importance of robust modeling with limited features. Techniques such as privacy-preserving aggregation, feature hashing, and differential privacy can be relevant, but they can reduce signal-to-noise ratio and complicate incident investigation. Platforms therefore face a tradeoff between model performance, auditability, and privacy guarantees. A consistent theme is that the detection problem cannot be separated from the data governance problem; model quality depends on the platform's ability to collect, retain, and interpret signals in a way that remains legitimate under applicable rules.

4 Detection Models and Risk Scoring for Adaptive Adversaries

A detection system for peer-to-peer fraud and misrepresentation must map heterogeneous signals into decisions under uncertainty and asymmetric costs [19]. The output is rarely a binary classification used in isolation; more often it is a calibrated risk score that drives a set of interventions with different user experience impacts. The technical objective is therefore better described as minimizing expected harm subject to constraints on friction, fairness, and operational capacity, while maintaining robustness to label noise, drift, and strategic manipulation.

A convenient abstraction treats each transaction or account action as an instance with features and an unobserved latent intent. Let X denote observed features derived from account attributes, behavioral logs, network structure, and content signals, and let $Y \in \{0,1\}$ denote an event of interest such as a confirmed fraud outcome or a high-severity harm outcome. A model $s_\theta(X)$ produces a score that approximates $\Pr(Y = 1 \mid X)$ after calibration. Decisions are made by comparing $s_\theta(X)$ to thresholds that may depend on context such as category, value, or user tenure. A basic cost-sensitive formulation is

$$\min_{\theta,\tau} \quad \mathbb{E}\Big[C_{\mathrm{FN}}\mathbf{1}\{s_\theta(X) < \tau, Y = 1\} \; + \; C_{\mathrm{FP}}\mathbf{1}\{s_\theta(X) \geq \tau, Y = 0\}\Big] \tag{1}$$

$$\text{where} \quad C_{\mathrm{FN}} = C_{\mathrm{loss}} + C_{\mathrm{trust}} + C_{\mathrm{ops}}, \quad C_{\mathrm{FP}} = C_{\mathrm{friction}} + C_{\mathrm{attrition}} + C_{\mathrm{appeal}}. \tag{2}$$

Here C_{FN} captures financial loss, trust damage, and operational downstream cost when harm occurs, while C_{FP} captures the cost of incorrectly restricting legitimate users, including appeal processing and potential churn. In practice these costs vary by transaction value and by user segment, which motivates context-dependent thresholds $\tau(v,u)$ or decision policies that select among multiple actions rather than a single block/allow outcome.

Because labels are noisy and incomplete, estimating $\Pr(Y = 1 \mid X)$ directly from observed labels can be misleading. A positive-unlabeled perspective treats observed positives as a subset of true positives with detection probability $\pi(X)$ [20]. If $\tilde{Y}$ is the observed label and Y is the true event, then $\Pr(\tilde{Y} = 1 \mid Y = 1, X) = \pi(X)$ and $\Pr(\tilde{Y} = 1 \mid Y = 0, X) = 0$ under a simplified assumption. The observed conditional probability satisfies $\Pr(\tilde{Y} = 1 \mid X) = \pi(X)\Pr(Y = 1 \mid X)$, so naive supervised learning conflates propensity to be detected with propensity to be fraudulent. Approaches to mitigate this include estimating $\pi(X)$ via audit sampling, using semi-supervised anomaly detection to model the dominant benign distribution, or using weak supervision where multiple heuristics and reviewer signals are combined into probabilistic labels.

Anomaly detection can be attractive early in the lifecycle when little labeled data exists for new attack patterns. Methods such as one-class classification, density estimation in embedding space, or reconstruction error from autoencoders can flag deviations from learned benign behavior. However, peer-to-peer platforms are naturally heterogeneous, and benign novelty is common, especially for new users and rare categories. Pure anomaly scores tend to correlate with newness and sparsity, creating high false positive rates. A more robust approach conditions on context and uses hierarchical models that pool strength across similar categories while allowing local variation. For example, a Bayesian model can represent transaction rates and dispute rates with partial pooling, allowing the platform to distinguish an unusually high dispute rate for a seller relative to peers in the same category and price band [21].

Sequence modeling is often more informative than static features because fraudulent intent manifests through temporal patterns. Let $(X_t)_{t=1}^{T}$ denote a sequence of events for an account, such as logins, listings, messages, payment attempts, and disputes. A hidden-state formulation can model latent risk Z_t that evolves over time and generates observed events. A simple state-space model is

$$Z_t = \alpha Z_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \tag{3}$$

$$X_t \mid Z_t \sim p(\cdot \mid Z_t), \tag{4}$$

where Z_t summarizes risk propensity and $p(\cdot \mid Z_t)$ is an emission distribution that can be implemented via a neural sequence model. The posterior $\Pr(Z_t \mid X_{1:t})$ provides a principled way to update risk as new evidence arrives. In operational terms, this supports interventions that escalate with accumulating evidence rather than reacting to single spikes that may be benign.

Graph-based modeling addresses collusion and shared infrastructure. Construct a heterogeneous graph with nodes representing accounts, devices, payment tokens, addresses, listings, and transactions, and edges representing observed relationships such as "used device," "paid with token," "messaged," or "transacted." Fraud risk then becomes a problem of inference over relational structure. Message passing models can learn representations that propagate risk through the network while attenuating noise. A generic update for node representation $h_v^{(k)}$ at layer k is

$$m_v^{(k)} = \sum_{u \in \mathcal{N}(v)} \phi^{(k)}\big(h_v^{(k)}, h_u^{(k)}, e_{uv}\big), \tag{5}$$

$$h_v^{(k+1)} = \sigma\big(W^{(k)} h_v^{(k)} + U^{(k)} m_v^{(k)}\big), \tag{6}$$

where $\mathcal{N}(v)$ is the neighborhood of v, e_{uv} encodes edge type and attributes, $\phi^{(k)}$ is an aggregation function, and σ is a nonlinearity. Such models can detect patterns like many accounts connected to a small set of devices or payment instruments, or tightly knit clusters with unusually reciprocal interactions [22]. The risk is that shared legitimate infrastructure creates false associations, so graph models must be regularized and evaluated under realistic household and organizational sharing patterns.

A complementary approach uses probabilistic graphical models or belief propagation when interpretability and uncertainty quantification are priorities. For instance, if a device is shared by multiple accounts, the platform may maintain a posterior probability that the device is compromised and update account risks accordingly. This style can be easier to audit but may be less flexible in capturing complex behavior than learned embeddings. Hybrid systems can combine learned representations with probabilistic calibration layers to provide both performance and uncertainty estimates.

Calibration is central because operational decisions depend on score meaning. A model that ranks well but is poorly calibrated can cause unstable policy outcomes when thresholds shift [23]. Calibration methods such as isotonic regression or temperature scaling can map raw scores to probabilities, but calibration itself can drift as attacker behavior changes. A practical deployment strategy monitors calibration on recent cohorts and recalibrates periodically, while ensuring that recalibration does not inadvertently weaken defenses for emerging threats. Uncertainty estimation, through ensembling or Bayesian approximations, can support triage by routing high-uncertainty cases to manual review while automating clear cases.

Cost-sensitive learning extends beyond thresholding because intervention costs vary. If an action is a step-up verification rather than a block, the false positive cost may be lower but still nontrivial due to abandonment. Decision-focused modeling can directly optimize expected utility of actions. Let $a \in \mathcal{A}$ denote an action set such as allow, hold funds, require verification, limit messaging, or block. If $L(a, Y, X)$ is the loss of taking action a when outcome is Y under context

X, then an optimal policy under a probabilistic model chooses [24]

$$a^*(X) = \arg\min_{a \in \mathcal{A}} \quad \mathbb{E}\big[L(a, Y, X) \mid X\big]. \tag{7}$$

This formulation encourages modeling the conditional distribution of outcomes relevant to each action, not merely a single fraud label. For example, holding funds may be effective for chargeback-prone transactions but less relevant for misrepresentation disputes that hinge on item condition.

Human-in-the-loop review introduces additional structure. Reviewers produce labels, but they also take actions and gather evidence. If reviewer capacity is limited, the platform must select which cases to review. A triage policy can be framed as maximizing expected value of information under a budget constraint. If $\Delta(X)$ is the expected reduction in loss from reviewing a case with features X, and $c(X)$ is review cost, then selection aims to maximize $\sum \Delta(X_i)$ subject to $\sum c(X_i) \leq B$. Estimating $\Delta(X)$ requires modeling how reviewer outcomes alter future decisions and how often reviewer decisions are correct, which can vary by case type [25]. Over time, reviewer consistency and feedback loops become critical; inconsistent adjudication produces label noise that degrades models and increases appeals.

Adversarial robustness is a persistent concern. Attackers can attempt to evade models by modifying features, for example by changing text templates, varying transaction timing, or using distributed infrastructure. Unlike image adversarial examples, platform evasion often involves real operational constraints. Robustness can therefore be improved by emphasizing features that are costly for attackers to manipulate at scale, such as long-term behavioral consistency, cross-entity link structure, and settlement outcomes. Nonetheless, feature leakage can occur when enforcement messages reveal which behaviors triggered intervention. A governance-aware approach designs user-facing communications that support transparency and appeals without exposing high-leverage decision rules [26].

Drift detection and online learning are required because attacker tactics evolve. Monitoring can track distribution shifts in key features, changes in score distributions, and changes in outcome rates conditional on score. A stable system also needs counterfactual evaluation: if a model becomes more conservative and blocks more transactions, observed fraud rates may decrease even if underlying attempted fraud increases. Disentangling these requires randomized experiments or quasi-experimental designs where certain interventions are applied randomly to a subset of traffic to estimate causal effects. For instance, randomized step-up verification on a small share of medium-risk transactions can estimate how much fraud is deterred versus how much legitimate trade is lost to abandonment.

Fairness and proportionality constraints complicate optimization. If certain user groups are disproportionately affected by false positives, the platform may impose constraints on disparity in intervention rates or on error rates [27]. Let A denote a sensitive attribute or a proxy group defined for auditing. A constraint might limit the difference in intervention probability across groups:

$$\Pr(s_\theta(X) \geq \tau \mid A = a) - \Pr(s_\theta(X) \geq \tau \mid A = b) \leq \epsilon \quad \text{for audited groups } a, b. \tag{8}$$

In practice, platforms may not collect sensitive attributes, and group definitions may rely on imperfect proxies, so fairness auditing becomes statistically and ethically complex. Additionally, naive parity constraints can reduce protection for groups that are more frequently targeted by fraud. A more operational approach focuses on measuring harm and ensuring that protective interventions do not systematically exclude legitimate participants, while maintaining strong safeguards against discriminatory feature usage.

Misrepresentation detection specifically often benefits from multimodal modeling. Listing text, images, and metadata can be embedded into representations that capture semantic consistency, such as whether claimed brand names align with visual cues, or whether description patterns match known scam templates [28]. However, definitive authenticity inference is rarely possible. Models can instead estimate the probability of post-transaction disputes related to authenticity or condition, enabling pre-transaction warnings, escrow holds, or seller verification for high-risk combinations. This shifts the objective from proving wrongdoing to reducing expected harm.

Finally, evaluation metrics must align with deployment. AUC can be misleading when base rates are low and when costs are asymmetric. Precision at operationally relevant review capacity, expected loss reduction, and calibration error are often more meaningful. Because interventions change behavior, offline evaluation must be complemented with online monitoring and experiments [29]. A detection model that achieves a low false positive rate such as 0.5% can still generate large absolute volumes of impacted users at scale, making careful threshold governance and appeal workflows necessary for sustainability.

5 Transaction Security Architectures as Complementary Controls

Detection models infer risk, but transaction security architectures shape incentives and constrain feasible attacks. In peer-to-peer trading, security is not only cryptographic confidentiality or integrity; it is the design of protocols and workflows that reduce ambiguity, increase evidentiary quality, and allocate liability in ways that discourage misconduct. Many losses occur not because a platform lacks predictive signal, but because the transaction protocol allows disputes to remain ambiguous or allows irreversible value transfer before verification.

Escrow-like mechanisms are a foundational control. When funds are held until fulfillment conditions are met, the buyer's risk of non-delivery decreases and the seller's risk of frivolous refund claims can be managed through evidence requirements and time windows. The design details matter: what constitutes fulfillment, how long funds are held, what evidence is required, and how exceptions are handled [30]. For shipped goods, integration with carrier tracking can provide signals, but tracking is imperfect and can be manipulated in some contexts. For local exchanges, escrow can be coupled with check-in confirmations or optional identity verification at pickup locations, but these increase friction. A platform can use risk-based holds, applying longer holds or additional verification only for high-risk transactions, thereby balancing safety and usability.

Authenticated messaging and controlled communication channels improve security by reducing off-platform redirection and preserving evidence for dispute resolution. When communication occurs within platform channels, the platform can enforce policies against sharing sensitive payment details and can provide structured prompts that encourage safe behavior, such as reminding users not to complete payments outside the platform. Evidence preservation is especially important for misrepresentation disputes, where claims hinge on what was promised. Protocol design can encourage structured listing attributes and standardized condition descriptions, reducing semantic ambiguity [31]. However, overly rigid schemas can exclude niche items and can incentivize users to misclassify items to fit available categories, so schemas must be flexible enough to represent real variability.

Identity and account integrity controls reduce repeated abuse. Multi-factor authentication, device binding, and account recovery safeguards can reduce account takeover and unauthorized use. Strong identity checks can deter serial fraudsters but can create barriers for legitimate users without traditional documentation, and they can increase the consequences of data breaches. Risk-based identity verification can mitigate these concerns by focusing on actions that increase exposure, such as high-value listings, frequent high-velocity sales, or rapid cross-border activity. Importantly, identity verification should not be treated as a one-time event; ongoing account integrity monitoring can detect changes in device, network, and behavior that suggest compromise.

Payment rail choices and liability allocation have first-order effects [32]. If the platform bears chargeback liability, it has stronger incentives to screen payments and reduce fraud, but it also may impose stricter controls that affect legitimate users. If users bear liability, platform trust may suffer and transaction volume may decline. Hybrid approaches include seller protection programs contingent on following platform protocols, and buyer protection programs contingent on paying through platform rails. These programs should be designed to avoid moral hazard, where parties behave less cautiously because they expect reimbursement. One approach is to incorporate deductibles or to limit protection for repeated claims without strong evidence, though this must be balanced against the possibility that some users are repeatedly targeted.

Evidence capture can be integrated into the protocol. For example, requiring photo evidence at packaging or at delivery can deter certain forms of misrepresentation and provide data for adjudication, but it can be burdensome and privacy-sensitive [33]. For high-risk categories, optional third-party verification or authenticity checks can reduce disputes, but they increase cost and can create new points of failure. A platform can instead encourage voluntary verification through incentives and transparently labeled trust tiers, while keeping baseline access open. The security architecture should also consider how evidence is stored, who can access it, and how long it is retained, as these choices interact with privacy obligations.

Reputation systems should be designed to be resistant to manipulation. Weighting by transaction value, tenure, and diversity of counterparties can reduce the impact of low-value self-dealing, but it can also disadvantage new legitimate users. Incorporating dispute outcomes and complaint rates can improve signal, yet it risks embedding policy bias. A useful design principle is to separate public reputation displays, which should be conservative and resistant to gaming, from internal risk scores, which can incorporate high-dimensional signals not suitable for public exposure [34]. Internal scores can be updated quickly and can guide risk-based protocol adjustments without directly incentivizing attackers to optimize visible metrics.

Dispute resolution protocols are a security mechanism because they shape incentives and expected payoffs. Clear timelines, standardized evidence requirements, and proportional remedies reduce the ability to exploit ambiguity. For misrepresentation, remedies such as partial refunds or return requirements can reduce incentive for opportunistic claims. For non-delivery, requiring carrier-confirmed delivery or verified pickup can reduce ambiguity. However, strict evidence requirements can disadvantage users with limited capacity to produce documentation, so protocols should include accessible paths and support. Automated dispute triage can speed resolution but must be monitored for systematic errors and must allow escalation [35].

System design can incorporate throttling and rate limits to reduce rapid exploitation. Limits on listing velocity, messaging volume, and payment attempts can reduce the speed at which attackers extract value. The challenge is that high-activity legitimate sellers can be impacted. Risk-based throttling uses model outputs to adapt limits, applying stricter constraints only when risk is elevated. This requires careful calibration and monitoring because throttling can change observed behavior, potentially creating self-reinforcing signals if the model interprets throttled behavior as suspicious.

Cryptographic mechanisms can complement but rarely replace broader protocol design. Digital signatures can ensure message integrity and reduce repudiation, and secure logging can improve auditability [36]. For certain digital goods or tokenized assets, atomic settlement protocols can reduce delivery ambiguity. However, most peer-to-peer trading involves physical goods where cryptography cannot verify item condition. The primary role of cryptographic controls is therefore to protect account integrity, secure communications, and ensure audit trails for actions that have financial consequences.

A system-level view treats transaction security and detection as a coupled design problem. Detection models prioritize where friction is added, while security protocols determine what evidence and outcomes are available for future learning. By designing protocols that reduce ambiguity and standardize evidence, platforms improve both immediate loss prevention and long-term model quality. Conversely, protocols that externalize risk to users or that leave key events unobserved create blind spots that even sophisticated models cannot reliably fill [37].

6 Policy Implications, Governance, and Economic Incentives

Policy and governance choices influence fraud prevalence by shaping incentives for platforms, buyers, and sellers, and by constraining the feasible design space for data collection and enforcement. In peer-to-peer platforms, the platform is often the only scalable institution capable of providing enforcement at the point of interaction. Yet the platform's incentives are mixed:

reducing fraud protects trust and lowers operational losses, but increasing friction can reduce growth and liquidity. The policy environment, including consumer protection, payment regulation, privacy law, and automated decision-making oversight, changes the platform's objective function and the mechanisms available for risk management.

A central governance question is liability allocation. If platforms are held responsible for counterfeit goods, unauthorized payments, or unsafe transactions, they may invest more in verification and security, but they may also restrict participation. Conversely, minimal liability can reduce platform investment and shift losses to users [38]. An economically grounded approach views platform policy as choosing a level of verification and enforcement that balances deterrence against market access. The equilibrium can be modeled as a strategic interaction between the platform and adversaries. Let the platform choose enforcement intensity $p \in [0, 1]$ representing the probability that a harmful attempt is detected and sanctioned, and let attackers choose attempt intensity $q \geq 0$ representing the rate of harmful attempts. If the attacker's expected payoff per attempt is V and the expected penalty upon detection is F, then an attacker's expected utility can be represented as

$$U_{\text{att}}(q, p) = q\big((1-p)V - pF\big) - K(q), \tag{9}$$

where $K(q)$ is a convex cost capturing operational constraints and risk. The platform's expected loss can be represented as

$$U_{\text{plat}}(p, q) = -q(1-p)L - C(p), \tag{10}$$

where L is expected loss per successful attempt and $C(p)$ is the cost of enforcement, including friction and operational expense [39]. In such a model, increasing p reduces successful attacks but can increase $C(p)$. Policy interventions that increase F through legal penalties or improved identity binding can reduce attack incentives without requiring the platform to increase p as much, though such interventions may raise privacy and access concerns. While simplified, this framing highlights why protocol design and legal enforcement can be complements to detection: they change the attacker's payoff structure.

Transparency and due process are increasingly salient. Automated restrictions, fund holds, and account suspensions can be disruptive, particularly for users who depend on platforms for income. Governance norms therefore push toward explainability, appeal mechanisms, and proportional sanctions. However, full transparency can reveal detection features and enable evasion [40]. A practical balance is to provide users with actionable, policy-based explanations that describe the violated rule category and the required remediation steps, without revealing high-dimensional signals or precise thresholds. Appeals processes should be designed to correct errors efficiently and to feed back into model improvement, but they must be protected against being used as reconnaissance by adversaries. Rate-limiting appeals, requiring minimal evidence, and using specialized reviewer teams for high-impact cases are operational approaches, but they must be implemented carefully to avoid creating barriers for legitimate users.

Privacy regulation and user expectations constrain data usage. Device and network metadata, message content, and identity documents are sensitive. Policies may require data minimization, purpose limitation, and retention limits. These constraints can reduce model performance if not addressed through better measurement design [41]. One approach is to use derived features and embeddings, strict access controls, and privacy-preserving aggregation, while reserving raw data access for narrow security investigations. Governance should also address how data is shared with third parties, including payment processors and law enforcement, and how cross-border transfers are handled.

Fairness considerations arise because risk signals can correlate with socioeconomic factors, geography, and language. If models use proxies such as neighborhood-level signals or device types, they may inadvertently concentrate enforcement on certain groups. Even when sensitive attributes are not collected, disparate impact can occur. Governance therefore requires regular auditing using available proxies, careful review of feature sets, and monitoring of error rates and appeal outcomes. Importantly, fairness should be evaluated not only in terms of intervention rates but also in terms of protection: under-enforcement in communities that are more frequently targeted can increase harm [42]. A harm-centered audit asks whether the system reduces victimization equitably, while keeping false positives within acceptable bounds.

Cross-jurisdiction issues complicate policy. A transaction may occur between parties in different legal regimes, with different consumer protection rules and enforcement capabilities. Platforms often respond by standardizing policies globally to simplify operations, but this can conflict with local requirements. Alternatively, platforms can localize rules and remedies, but then attackers can exploit jurisdictional differences. A risk-based approach may apply stricter controls to cross-border transactions or to categories associated with high dispute ambiguity, but this can reduce legitimate international trade. Policymakers and platforms face a tradeoff between openness and safety that is not purely technical [**?**].

Regulation related to financial crime can intersect with peer-to-peer trading when platforms facilitate payments or store value. Requirements for customer due diligence, transaction monitoring, and reporting can increase detection capacity but can also raise compliance cost and create exclusion risks for users lacking documentation. In peer-to-peer marketplaces that are not primarily financial institutions, the applicability and scope of such obligations can be contested, leading to uneven compliance across platforms. From a system design perspective, the most robust posture is to design monitoring and identity processes that are modular and scalable, allowing the platform to adapt to evolving obligations without rebuilding the entire detection stack.

Platform governance also includes internal accountability. Detection models influence user outcomes, so they require documentation of training data, label definitions, known failure modes, and monitoring plans. Model changes should be evaluated not only for predictive metrics but also for downstream impacts such as appeal rates, user churn, and category-level liquidity [43]. Policy changes such as altering refund rules or verification requirements should be assessed for their causal effect on fraud and on legitimate trade. Randomized experiments can provide evidence, but governance must ensure experiments do not expose users to undue harm. When experiments are infeasible, quasi-experimental methods and careful observational analysis become important, though they carry more uncertainty.

Economic incentives can be shaped by platform fee structures and protections. If seller fees are high, sellers may seek off-platform payment, increasing risk. If buyer protection is generous with low evidentiary burden, opportunistic claims may rise. If seller protection is generous, buyers may face counterfeit risk [44]. A balanced design aligns incentives by rewarding compliance with secure protocols, such as shipping with trackable carriers, using platform payments, and providing accurate listings. Trust tiers can provide benefits such as faster payouts or higher visibility for users with strong histories, but they must be robust to reputation laundering. Policy governance should also consider how promotions and subsidies can be abused and should incorporate anti-abuse constraints into marketing design.

Finally, policy implications extend to information sharing and industry coordination. Fraud often spans multiple platforms, and shared threat intelligence can improve defenses. Yet sharing must respect privacy and competition constraints. Privacy-preserving sharing mechanisms, such as sharing hashed indicators or aggregated patterns, can help, but they can also create false matches and accountability issues [45]. Governance frameworks that define what is shared, how it is validated, and how errors are corrected are necessary to avoid amplifying false accusations.

Overall, policy and governance are not external constraints applied after technical design; they are part of the system that determines what signals exist, what actions are feasible, and what tradeoffs are acceptable. Platforms that integrate detection, protocol design, and governance can reduce fraud and misrepresentation while maintaining market access, but they must continuously adapt as adversaries and regulatory environments evolve.

7 Conclusion

Fraud, misrepresentation, and transaction insecurity in peer-to-peer trading platforms arise from a combination of information asymmetry, weak enforcement, and adaptive adversaries operating in software-mediated markets. Technical detection models are an important component of mitigation, but their effectiveness depends on measurement quality, lifecycle-aware modeling, and the surrounding transaction protocols that determine what evidence and outcomes are observable. A system-level approach treats fraud risk as sequential and partially observed, acknowledges that labels are delayed and biased by interventions, and designs models that output calibrated risk scores suitable for differentiated actions rather than binary outcomes alone.

Detection modeling in this setting benefits from combining supervised learning with approaches robust to incomplete labels, including semi-supervised and anomaly-oriented methods, graph-based inference for collusion and shared infrastructure, and sequence models that capture temporal intent signals [46]. Cost-sensitive objectives aligned to operational harms help connect modeling to platform decisions, while calibration and uncertainty estimation improve threshold governance and triage. Because interventions change observed data, causal evaluation and monitoring under drift are essential for maintaining validity over time. Human review remains important for ambiguous cases and for generating high-quality labels, but it must be integrated as part of a feedback-aware learning system rather than treated as an external correction.

Transaction security architectures complement detection by reshaping incentives and reducing ambiguity. Escrow-like holds, authenticated messaging, structured evidence capture, and risk-based throttling reduce the profitability of common fraud strategies and improve dispute resolution reliability. These controls also improve the learning environment by standardizing signals and outcomes, enabling more stable calibration and better attribution of harm causes. Security design choices, however, introduce friction and privacy sensitivity, requiring careful balancing through context-dependent controls [47].

Policy and governance determine the acceptable trade space for data usage, transparency, and enforcement. Liability allocation, privacy constraints, and automated decision-making oversight influence which signals can be collected and how decisions must be explained and appealed. Fairness auditing and proportional sanctions are necessary to reduce the burden of false positives while ensuring that protection is not unevenly distributed. Cross-jurisdiction complexity and financial crime obligations further complicate design, motivating modular controls and evidence-based policy iteration.

A practical implication is that platforms should evaluate risk management as an integrated system: measurement pipelines, models, protocols, and governance mechanisms must be co-designed and co-monitored. Improvements in one component can be negated by weaknesses in another, while coordinated changes can yield nonlinear gains by simultaneously deterring adversaries and improving inference. Continued progress will likely depend on better handling of feedback-induced bias, stronger lifecycle-specific modeling, and governance processes that provide transparency and recourse without enabling evasion, all while preserving legitimate participation in peer-to-peer markets [48].

References

[1] S. S. M. Bala and S. Kumari, “Comprehensive analysis of variants of tf-idf applied on lda and lsa topic modelling,” *International Journal of Engineering and Advanced Technology*, vol. 9, pp. 531–535, 8 2020.

[2] S. M. Rahman, “Consumer expectation from online retailers in developing e-commerce market: An investigation of generation y in bangladesh,” *International Business Research*, vol. 8, pp. 121–137, 6 2015.

[3] N. L. Gutiérrez, “Comparative analysis of online and physical consumer behavior towards csr,” 12 2016.

[4] N. V. Patel, "Selection bias in two-sided e-commerce marketplaces: A framework for propensity score matching implementation," *Journal of Business Intelligence Systems and Computational Social Science Applications*, vol. 11, no. 5, pp. 1–20, 2021.

[5] J. Hu, "Money creation in big tech lending," *SSRN Electronic Journal*, 1 2022.

[6] B. Jullien and I.-U. Park, "Seller reputation and trust in pre-trade communication," 10 2009.

[7] A. R. Reuber and E. Fischer, "Signalling reputation in international online markets," *Strategic Entrepreneurship Journal*, vol. 3, pp. 369–386, 12 2009.

[8] S. Keskin, "Covid-19 pandemi sürecinde online market alışverişine yönelik tüketici davranışında cinsiyete göre İzlenen değişimin analizi: Ankara Çalışması," 7 2021.

[9] M. Arefi and A. M. Amini, "Investigating intermediaries that participate in internet commerce to determine the elements of online price dispersion," *International Journal for Digital Society*, vol. 1, pp. 76–85, 6 2010.

[10] M. Norris and S. West, *eBusiness Essentials: Technology and Network Requirements for Mobile and Online Markets, Second Edition - Putting the 'e' into Your Business*. Wiley, 10 2001.

[11] R. D. Moss, "Civil rights enforcement in the era of big data: Algorithmic discrimination and the computer fraud and abuse act," 3 2016.

[12] D. Masclet and T. Pénard, "Pourquoi évaluer son partenaire lors d'une transaction à la ebay ? une approche expérimentale," 6 2006.

[13] F. Ghorbanian and M. Jalali, *ICCAI - A Novel Hybrid Algorithm for Sentiment Analysis via Classifier Ensembles for Online Shops User Using User Generated Contents and Review*. ACM, 4 2020.

[14] U. N. Haque and R. Mazumder, "A study on the relationship between customer loyalty and customer trust in online shopping," *International Journal of Online Marketing*, vol. 10, pp. 1–16, 4 2020.

[15] E. Muntoni, "Travel distribution marketing: how the internet has influenced traveller's buying behaviour in different european countries," 3 2013.

[16] M. P. Banu and A. M. Begum, "A study on consumer behaviour towards online shopping in tiruchirappalli," *INTERNATIONAL JOURNAL OF MANAGEMENT AND SOCIAL SCIENCES*, vol. 8, pp. 81–84, 9 2018.

[17] A. J. Larson and D. Sachau, "Effects of incentives and the big five personality dimensions on internet panellists' ratings:," *International Journal of Market Research*, vol. 51, pp. 687–706, 1 2009.

[18] Y. Jiang, H. H. Wang, S. Jin, and M. S. Delgado, "The promising effect of a green food label in the new online market," *Sustainability*, vol. 11, pp. 796–, 2 2019.

[19] A. Sasso, M. Hernández-Alava, J. Holmes, M. Field, C. Angus, and P. Meier, "Strategies to cut down drinking, alcohol consumption, and usual drinking frequency: Evidence from a british online market research survey.," *Social science & medicine (1982)*, vol. 310, pp. 115280–115280, 8 2022.

[20] N. Ito, S. Inoue, T. Higuchi, H. Kobayashi, R. Mori, and T. Ishida, "Consumers' attitudes toward online food purchases in china: Segmentation analysis of online food market," *Japanese Journal of Agricultural Economics*, vol. 24, pp. 40–45, 3 2022.

[21] S. Feuß, M. Brettel, and H.-W. Schroiff, "Understanding the effect of sustainable products on consumer purchase and return behavior - an empirical analysis in online markets," 7 2020.

[22] M. Hardey, “Conference notes - the social context of online market research: an introduction to the sociability of social media,” *International Journal of Market Research*, vol. 51, pp. 562–564, 7 2009.

[23] M. Laouenan, morgane laouenan, P. Deschamps, G. Chapelle, and X. Lambin, “Discrimination on online markets: Evidence from a field experiment,” 3 2021.

[24] A. Ahamed, “As an online market place, the customer servicer acivities & customer satisfaction of kaymu.com.bd,” 5 2015.

[25] N. V. Patel, “Applying synthetic control methods to address causal identification challenges in the ride-hailing industry,” *Journal of Data Science, Predictive Analytics, and Big Data Applications*, vol. 8, no. 7, pp. 27–49, 2023.

[26] X. Zhao, K. Zhao, and J. Deng, “Geography still matters: Examine the role of location in online markets for foreign branded products,” *Decision Sciences*, vol. 50, pp. 285–310, 9 2018.

[27] M. Johari and A. D. Laksito, “The hybrid recommender system of the indonesian online market products using imdb weight rating and tf-idf,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, pp. 977–983, 10 2021.

[28] W. Li, D. Wu, and H. Xu, “Reputation in china’s online auction market: Evidence from taobao.com,” *Frontiers of Business Research in China*, vol. 2, pp. 323–338, 7 2008.

[29] X. Liu, E. Li, and Q. Li, *WHICEB - The Impact of Online Store Characteristics on Service Recovery Satisfaction in C2C Online Markets*. 6 2014.

[30] A. Kigerl, “Behind the scenes of the underworld: Hierarchical clustering of two leaked carding forum databases:,” *Social Science Computer Review*, vol. 40, pp. 089443932092473–640, 6 2020.

[31] S. Basak, I. S. Swazan, and D. Das, “Gen z’s intention to return product online: A regression analysis of young indian consumers,” in *Breaking Boundaries*, Iowa State University Digital Press, 9 2022.

[32] *MIPRO - The role of perceived privacy and perceived security in online market*, 5 2012.

[33] M. Ruslan, S. Boguslaw, and Z. Patrycja, “E-commerce in the era of globalization,” *Herald of Kyiv National University of Trade and Economics*, vol. 136, pp. 67–78, 4 2021.

[34] P. Engström and E. Forsell, “Demand effects of consumers’ stated and revealed preferences,” 4 2013.

[35] D. D. Zeng, J. C. Cox, and M. Dror, “Coordination of purchasing and bidding activities across posted offer and auction markets,” *Information Systems and e-Business Management*, vol. 5, pp. 25–46, 8 2006.

[36] Q. Gu, X. Yang, and B. Liu, “Pricing decisions on online channel entry for complementary products in a dominant retailer supply chain,” *Sustainability*, vol. 12, pp. 5007–, 6 2020.

[37] U. Khabibah, T. I. Wardani, J. D. Pribadi, Y. Afandi, and Y. S. Oktora, “Pelatihan pemasaran online guna memperluas jaringan pemasaran produk ukm di lingkungan ibu-ibu rw 08 perumahan joyogrand merjosari malang,” 10 2020.

[38] P. Adela and T. Naiana, “Study regarding the obstacles which encumber the online selling increase,” 5 2009.

[39] *An Empirical Study on the Culture Products Price Dispersion in Online Shopping Market*, 3 2011.

[40] L. N. T. Quang, U. N. T. Thao, A. D. V. Nhi, and H. N. T. Phuong, “E-commerce price suggestion algorithm – a machine learning application,” in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 2598–2609, IEOM Society International, 8 2022.

[41] Y. Zabyelina and S. Noguera, *Organized Crime and the Pharmaceutical Industry*, pp. 224–241. Routledge, 6 2022.

[42] K. Sommermeyer and S. Pookulangara, "Liketoknow.it: The influencer economy, an exploratory study," 1 2019.

[43] U. Radkevitch, E. van Heck, and O. Koppius, "Buyer commitment and opportunism in the online market for it services," 8 2006.

[44] A. Popescu, D. Popescu, and C. State, *Hashtag Progress: The Digital Fingerprint of Web 2.0 on Tourism and Hospitality Industry Management—A Case Study for Romania*, pp. 555–564. Springer International Publishing, 3 2016.

[45] N. V. Patel, "Adaptive experimentation in marketplaces: Balancing exploration and revenue optimization," *Frontiers in Health Informatics*, vol. 11, pp. 760–786, 2022.

[46] M. E. E. Carmona, "La repercusión del e-commerce y el efecto de los marketplaces: Tendencias futuras en el sector del juguete," 1 2018.

[47] P. Očko, "Definition and topical problems of the information economy," *Politická ekonomie*, vol. 53, pp. 383–404, 6 2005.

[48] I. R. Marbun and R. Somya, "Perancangan data warehouse untuk data transaksi penjualan menggunakan schema snowflake studi kasus : Online market dataset," 5 2021.