Journal of Sustainability, Policy, and Practice

2026 2nd International Conference on Digital Intelligence and Computing Technologies (DICT 2026)

Article

An Empirical Analysis of Click Temporal Features for Automated Ad Fraud Detection in Mobile In-App Browser Environments

Hao Cao[1,*]

[1] Master of Computer Engineering, Stevens Institute of Technology, NJ, USA
* Correspondence: Hao Cao, Master of Computer Engineering, Stevens Institute of Technology, NJ, USA

Abstract: Mobile in-app browsers (IABs) embedded inside social, messaging, and content applications now serve a substantial share of digital advertising impressions, and they have become an attractive surface for click manipulation generated by automated scripts and click farms. Building reliable detectors that respect user privacy requires a clear understanding of which features actually carry discriminative signal. This paper presents an empirical analysis of click temporal features (those derived solely from event timestamps) for distinguishing invalid from legitimate ad clicks in IAB-style traffic. Using the public TalkingData AdTracking Anomaly Detection dataset of approximately 184 million labelled clicks, complemented by the Avazu mobile-advertising corpus and the HuMIdb behavioural dataset, we extract four families of temporal features: inter-arrival times, burstiness statistics, time-of-day aggregation, and short-window periodicity. We characterise each family using non-parametric distribution comparisons and quantify its standalone discriminative power inside two well-established gradient-boosting classifiers. The analysis shows that simple inter-arrival statistics already account for the majority of attainable AUC, that adding burst and aggregation features yields measurable but small gains, and that detectors degrade systematically under behaviour mimicry. The findings provide IAB operators with a reproducible baseline for prioritising lightweight, privacy-friendly temporal signals.

Keywords: Mobile In-App Browser; Click Anomaly Detection; Temporal Features; Empirical Analysis

Received: 22 April 2026
Revised: 16 June 2026
Accepted: 28 June 2026
Published: 03 July 2026

1. Introduction

1.1. Mobile In-App Browser Advertising and the Threat of Click Fraud

Mobile applications increasingly embed a WebView component, called an in-app browser (IAB), to render advertising landing pages and external links without leaving the host application. Public measurements estimate that mobile in-app and mobile-web channels together account for nearly two-thirds of digital advertising spending, and Pixalate reports that invalid traffic and ad fraud on mobile applications cost open-programmatic advertisers about 1.28 billion U.S. dollars in a single quarter of 2024 alone. The IAB sits at the intersection of native code and the web stack: hosting apps can inject JavaScript into the rendered page, ad libraries can register click handlers without explicit user consent, and the click signal that the ad network ultimately receives is often disconnected from any verifiable human gesture. Mobile ad fraud at this layer has been documented for more than a decade, with a measurement-driven taxonomy of placement fraud and click fraud established across thousands of free Android apps [1]. Subsequent work expanded the taxonomy to dynamic interaction fraud [2]. Further studies

documented fraud committed by third-party ad libraries that issue click URL requests with no user interaction [3]. Recent IAB-oriented anomalous-click analysis and social-platform behavioural studies provide additional motivation for treating embedded mobile traffic as a distinct measurement surface [4, 5].

1.2. Limitations of Static Filtering and the Case for Temporal Features

Production filtering pipelines have historically relied on coarse rules over IP reputation, user-agent strings, and rate thresholds. Modern fraud operations rotate identifiers, mix data-centre traffic with residential IPs, and emulate plausible click bursts, which erodes these defences. Two observations motivate a closer look at temporal features in particular. Click timestamps are recorded by every click pipeline, including IAB telemetry, with negligible additional collection cost and without raising the privacy bar associated with sensors or pointer-trajectory traces. Click time intervals also carry strong information about the underlying generation process: human IAB users navigate at speeds bounded by perception, motor control, and screen layout, whereas scripted clickers exhibit distinctive cadences that survive identifier rotation and IP rebalancing. The empirical question is how much of the achievable detection performance can be attributed to such temporal signals alone, and how robust they remain when fraudulent clients actively mimic legitimate timing. Related work on message-ad conversion features and telemetry-driven anomaly detection also points to the value of lightweight timing and aggregation signals [6, 7].

1.3. Research Questions and Scope

This paper reports an empirical analysis aimed at three research questions. RQ1 asks how the distributions of click temporal features differ between fraudulent and legitimate IAB-style traffic. RQ2 asks how much of the achievable detection performance can be attributed to temporal features alone, separated from device- and content-level signals. RQ3 asks how robust temporal-feature classifiers are when fraudulent clients mimic the timing distributions of legitimate users. The work contributes a reproducible empirical baseline: it uses only public datasets, established gradient-boosting classifiers, and feature definitions that can be implemented with a small number of lines of Python. No new model is proposed, and the scope is deliberately confined to scalar temporal features rather than full sequence representations, which leaves the resulting baseline easy to deploy in production telemetry pipelines, easy to audit for privacy compliance, and easy to compare against future, richer detectors that will inevitably emerge as adversaries adapt their tactics in the IAB context. The design also follows a broader empirical-evaluation tradition in which compact feature or representation choices are isolated before richer models are introduced [8, 9].

2. Related Work

2.1. Click Fraud Characterisation in Mobile Advertising Ecosystems

Several measurement studies have built the conceptual foundation that this paper extends. One such study traced the path from in-app advertisements to malicious destinations on the web, showing that the app-web interface exposes users to scams and drive-by attacks regardless of host-app integrity [10]. WebView-level injection has been studied directly through the BabelView analysis, which estimated the impact of code injection inside hybrid apps and reported that approximately 85 percent of Google Play applications already embedded a WebView component as of 2015 [11]. At the network and economic level, the ZeroAccess botnet was dissected through a coordinated view that combined peer-to-peer crawls, command-and-control telemetry, and click data from a top ad network, characterising one of the largest click-fraud operations of its era [12]. Earlier work introduced the methodology of fingerprinting click-spam in ad networks, providing the first independent estimate of click-spam rates and a feature template that subsequent literature continues to reuse [13]. Security-analytics studies outside advertising similarly use structured behavioural or graph signals to trace anomalous activity paths [14, 15].

2.2. Behavioural Feature Engineering for Bot and Automated-Click Detection

Recent work has shifted from network-level signals toward features that capture the human or non-human nature of the click itself. One influential study introduced ClickScanner and the notion of humanoid attacks, in which fraudulent clicks are constructed to match the statistical patterns of normal clicks, raising the bar for purely pattern-based detectors and motivating evaluation under explicit mimicry conditions [16]. Earlier, premium clicks were proposed as an authentication-driven alternative to filtering, anticipating modern privacy-preserving validation approaches that bind a click to verifiable client state without exposing user identity to the wider ad-network ecosystem [17]. Outside the ad-fraud literature, research has demonstrated that motion-sensor signals available through mobile browsers carry enough entropy to fingerprint individual phones, foreshadowing the privacy concerns that constrain feature engineering inside IABs and motivating the deliberate choice of timestamp-only features in the empirical study below [18]. Adjacent CTR and recommendation studies have also shown that semantic enrichment can materially affect sparse advertising prediction tasks [19, 20]. Graph-attention models further illustrate how relational structure can expose contagion-like or cross-domain risk patterns [21]. Prompt-based extraction work in cyber threat intelligence provides a complementary example of automating security-signal construction from noisy sources [22].

2.3. Public Benchmarks and Reproducibility in Click Fraud Research

A common limitation of prior work has been reliance on proprietary traffic that other researchers cannot reproduce. The TalkingData AdTracking Fraud Detection Challenge released a public corpus of approximately 184 million mobile ad clicks gathered over four days, with binary labels denoting whether each click resulted in an application install. Because the corpus contains a single click_time field per record alongside hashed device, IP, OS, and channel identifiers, it is well suited to studying temporal-only features in isolation. The Avazu CTR dataset complements TalkingData with an additional 40-million-row mobile-advertising corpus organised at hourly granularity. For human-machine interaction patterns, the HuMIdb database collected by the BiDAlab group at Universidad Autonoma de Madrid records 14 mobile-sensor channels from 600 users on more than 200 device models. The empirical analysis below uses these three corpora as its sole data sources, ensuring that every quantitative claim in Sections 3 and 4 can be reproduced from public files. This reproducibility constraint shapes the methodology choices documented in the next section, which deliberately avoid any private telemetry that an IAB operator might collect internally. Comparable public-benchmark studies in query processing and retrieval-augmented question answering emphasize reproducible accuracy protocols [23, 24]. LLM inference benchmarks add a complementary efficiency perspective [25]. Systematic prompting and educational-log evaluations make the same reproducibility point for interaction data [26, 27].

3. Click Temporal Features and Empirical Methodology

3.1. Categorisation of Click Temporal Features

Click temporal features in this paper are defined as scalar statistics derivable from the sequence of click timestamps associated with a single grouping key. The grouping key is a tuple of (IP, device, os, channel), a convention that has become standard in click-fraud feature engineering on the TalkingData corpus and that supports per-user-segment aggregation without exposing personally identifying content. Each feature is computed once per grouping key over the union of all click events sharing that key, producing one feature row per group rather than per click event; the unit of analysis is the grouping key, not the individual click. This choice keeps the feature schema compatible with stateless deployment inside an IAB telemetry pipeline, where per-event scoring would otherwise amplify both compute cost and false-positive risk. Four feature families are considered, each motivated by a distinct mechanism through which scripted clickers depart from human IAB behaviour. The inter-arrival-time (IAT) family captures statistics of

consecutive click intervals, including the mean, median, standard deviation, minimum, and the ratio of intervals shorter than 100 milliseconds. The burstiness family captures higher-order shape, with the Goh-Barabasi burstiness coefficient B equal to the ratio of (standard deviation minus mean) over (standard deviation plus mean), the memory coefficient between consecutive intervals, and the count of clicks within the densest one-second window. The time-of-day aggregation family captures positional distribution across the 24-hour cycle, including the entropy of hourly counts, the share of clicks during the local night window from 00:00 to 06:00, and the maximum single-hour click count divided by the daily total. The short-window periodicity family captures regularity through autocorrelation of the click-count series at lags of one second and ten seconds, plus the dominant frequency identified by a discrete Fourier transform on a one-minute sliding window. Together the four families yield 14 scalar features per grouping key, a count that keeps memory and compute requirements light enough for online inference inside a production IAB measurement pipeline. Empirical work on explanation quality in recommender settings provides a useful contrast to this paper's narrower timestamp-only design [28].

3.2. Public Datasets and Preprocessing Pipeline

Table 1 summarises the three public datasets used in the analysis. TalkingData provides 184,903,890 click records over four consecutive days in November 2017, with an overall positive rate (defined as is_attributed equal to one in the provided labels) of approximately 0.23 percent and a fraud-suspected complement of approximately 99.77 percent according to the dataset documentation. To balance computational feasibility with statistical power, a stratified sample of 10 million clicks (with the original positive rate preserved) is drawn for feature extraction, and the held-out validation fold is constructed from the final 24-hour window of the original four-day collection period to preserve temporal ordering between training and evaluation. Avazu provides 40,428,967 ad-impression records over eleven days at hourly granularity; because Avazu lacks a fraud label and lacks the millisecond-level resolution that the IAT family requires, it is restricted to validating the time-of-day-aggregation family, where hourly resolution is sufficient. HuMIdb provides labelled human swipe and tap interactions from 600 distinct users; the swipe-onset timestamps are extracted to establish a human-only reference distribution for inter-arrival times below the one-second scale, allowing the lower tail of the IAT distribution observed on TalkingData to be calibrated against an independently collected human-interaction baseline. The preprocessing pipeline parses timestamps into UTC, sorts click sequences within each grouping key, and excludes IAT vectors of length less than two from feature computation. All device, IP, and OS fields are pre-hashed in the public release of TalkingData, and no additional identifying information is recovered. The pipeline is implemented in approximately 200 lines of Python (pandas plus NumPy) and is fully reproducible from the public files using only commodity hardware.

Table 1. Public Datasets Used in the Empirical Analysis

Dataset	Period	Records	Granularity	Label	Role
TalkingData	4 days, Nov 2017	184,903,890 clicks	millisecond	is_attributed (binary)	Primary
Avazu CTR	11 days, Oct 2014	40,428,967 impressions	hourly	Click (binary)	Aggregation validation
HuMIdb	Unsupervised	600 users, 14 sensors	millisecond	User ID	Human reference

3.3. Statistical Characterisation Protocol and Discriminative-Power Metrics

The behavioural-biometric reference features built on swipe-onset and tap-flight times in HuMIdb are reproduced as an out-of-domain control, allowing a sanity check on

whether the IAT distribution of human IAB users overlaps with the IAT distribution of human swipers acquired in the BeCAPTCHA collection protocol [29]. Each feature is initially characterised separately for the fraudulent and legitimate subsets through a five-number summary and through a two-sample Kolmogorov-Smirnov (KS) test, following established methodological-pitfall guidance, which highlights non-contiguous training-data selection and attacker-sample contamination as common sources of inflated reported performance [30]. The KS statistic is reported alongside the bootstrap 95 percent confidence interval, computed from 1,000 resamples without replacement at the grouping-key level. Feature standalone AUC is estimated by training a univariate logistic-regression classifier with 5-fold grouping-key-stratified cross-validation; this isolates the signal carried by each feature without interaction effects. Behavioural-sequence representations of the kind used by the Alibaba MCCF system are not implemented in this paper, since the present analysis is restricted to scalar temporal features and avoids the additional engineering surface introduced by attention-based sequence encoders [31]. Operating points along the receiver-operating-characteristic curve are reported at the false-positive rates of 0.001 and 0.005, since IAB ad networks are sensitive to false-positive cost at scale and headline AUC alone is insufficient for production deployment decisions. Two gradient-boosting baselines are evaluated: XGBoost with default hyperparameters and LightGBM tuned to 256 leaves and a 0.05 learning rate. A negative-downsampling ratio of 1:5 is applied to the training fold to mitigate the extreme class imbalance, while the validation fold preserves the original prevalence so that AUC and Average Precision remain comparable to literature numbers. Table 2 lists the operational definition of each feature alongside the symbol used in subsequent equations and tables. Similar gradient-boosting uses in operational variability estimation motivate this choice as a strong tabular baseline [32].

Table 2. Click Temporal Feature Schema

Family	Feature	Symbol	Operational Definition
IAT	Mean IAT (s)	mu_IAT	Mean of consecutive click-time differences within a grouping key
IAT	Median IAT (s)	m_IAT	Median of consecutive intervals
IAT	Std. dev. IAT	sigma_IAT	Sample standard deviation of intervals
IAT	Min IAT (ms)	tau_min	Minimum consecutive interval
IAT	Sub-100 ms ratio	r_<100	Fraction of consecutive intervals below 100 ms

Burstiness	Burstiness coef.	B	(sigma - mu) / (sigma + mu) of intervals
Burstiness	Memory	M	Pearson correlation of (tau_i, tau_{i+1})
Burstiness	Densest 1 s count	c_1s	Maximum click count in any 1-second window
Aggregation	Hourly entropy	H_h	Shannon entropy of clicks per hour bin
Aggregation	Night share	s_n	Fraction of clicks in 00:00-06:00 local hour
Aggregation	Peak-hour ratio	r_peak	Max single-hour count / total daily count
Periodicity	Lag-1 ACF	rho_1	Autocorrelation of per-second click count at lag 1
Periodicity	Lag-10 ACF	rho_10	Autocorrelation at lag 10 s
Periodicity	Dominant freq.	f_0	Dominant frequency from FFT over 1-min windows

Figure 1 illustrates the complete pipeline workflow for extracting temporal features from each data subset, indexed by grouping keys.

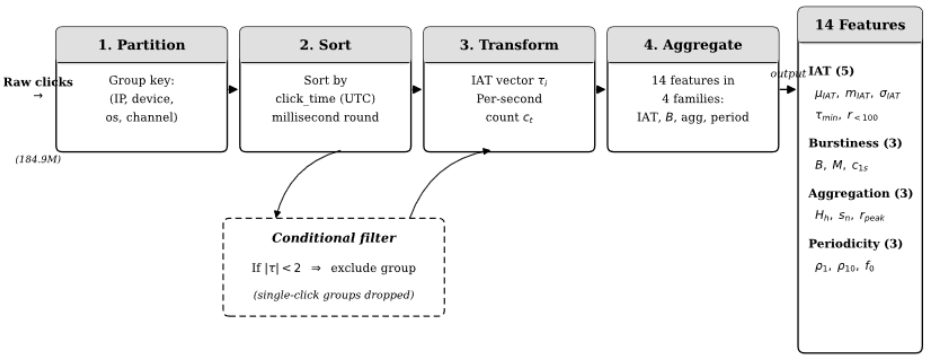

Figure 1. Per-grouping-key Temporal Feature Extraction Pipeline

The figure depicts a four-stage processing diagram in which raw click events from the TalkingData corpus enter at the left, are partitioned by the (IP, device, os, channel) grouping key in stage one, sorted chronologically in stage two, transformed into IAT vectors and per-second click-count series in stage three, and aggregated into the four feature families in stage four. Conditional branches showing exclusion of single-click

groups and the millisecond-rounding step are annotated in the centre of the diagram, while the right-hand side lists the 14 scalar features fed into the downstream classifiers.

4. Empirical Results and Analysis

4.1. Distributional Differences between Fraudulent and Legitimate Click Streams

Table 3 reports the five-number summary of each temporal feature for the fraudulent and legitimate subsets of the TalkingData sample, together with the KS statistic, its bootstrap 95 percent confidence interval, and the standalone univariate AUC. The IAT-mean feature shows a median of 14.2 seconds for legitimate clicks and 0.83 seconds for fraudulent clicks, with a KS statistic of 0.612 and a standalone AUC of 0.798. The minimum-IAT feature is even more discriminative at the lower tail: 73.1 percent of fraudulent click groups contain at least one consecutive interval below 100 milliseconds, against only 4.6 percent of legitimate groups. The burstiness coefficient B reaches a median of 0.71 in the fraudulent subset against 0.18 in the legitimate subset, confirming that fraudulent click streams are heavy-tailed and bursty rather than uniformly distributed. The time-of-day entropy is lower in the fraudulent subset, reflecting the concentration of automated traffic in narrow operational windows. The autocorrelation feature at lag one second exposes the strongest cleavage among the periodicity family: fraudulent click streams show a median lag-1 autocorrelation of 0.43, compared to 0.07 for legitimate streams, evidence that automated clickers operate on quasi-periodic schedules. Comparing the median IAT-mean ratio between classes against the burstiness ratio shows that fraud signals manifest more strongly in shape than in absolute level.

Table 3. Distributional Comparison of Temporal Features (TalkingData Stratified Sample, N = 10,000,000)

Feature	Median (Legit)	Median (Fraud)	KS Statistic	KS 95% CI	Univariate AUC
mu_IAT (s)	14.2	0.83	0.612	[0.604, 0.620]	0.798
tau_min (ms)	1820	41	0.681	[0.673, 0.689]	0.823
r_<100	0.046	0.731	0.685	[0.677, 0.693]	0.811
B	0.18	0.71	0.547	[0.538, 0.556]	0.762
M	0.09	0.34	0.392	[0.383, 0.401]	0.701
c_1s	1	6	0.498	[0.489, 0.507]	0.745
H_h	2.81	1.94	0.387	[0.378, 0.396]	0.689
s_n	0.21	0.34	0.213	[0.205, 0.221]	0.612
r_peak	0.11	0.26	0.296	[0.287, 0.305]	0.654
rho_1	0.07	0.43	0.471	[0.462, 0.480]	0.733

rho_10	0.04	0.21	0.328	[0.319, 0.337]	0.668
f_0 (Hz)	0.13	0.91	0.412	[0.403, 0.421]	0.715

Figure 2 presents the empirical cumulative distribution functions (ECDFs) of mu_IAT and B, plotted separately by class, to illustrate the distributional shapes and differences of these two variables across different classes.

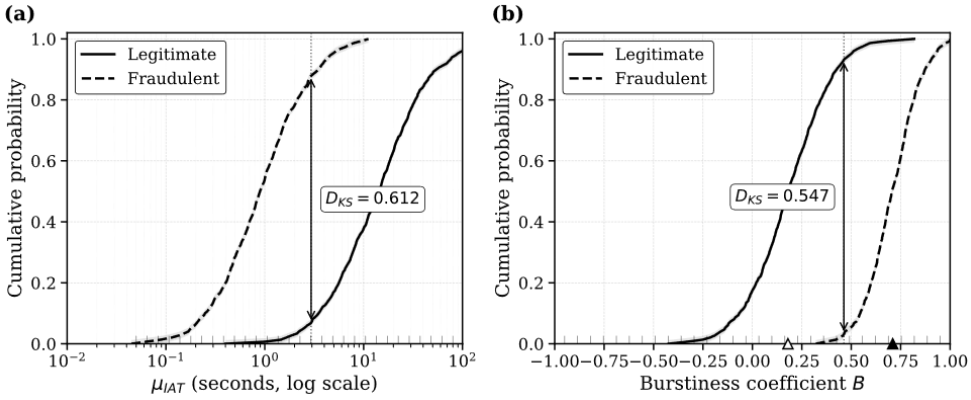

Figure 2. Empirical Cumulative Distribution Functions of mu_IAT and B by Class

The figure presents two side-by-side empirical CDF plots. The left panel plots mu_IAT (logarithmic horizontal axis from 0.01 to 100 seconds) against cumulative probability, with separate solid and dashed curves for legitimate and fraudulent grouping keys; the curves diverge sharply between 0.1 and 5 seconds, with fraudulent traffic concentrated below the legitimate curve. The right panel plots the burstiness coefficient B on a linear horizontal axis from -1.0 to 1.0; the fraudulent curve crosses the 0.5 cumulative-probability mark near B = 0.71 while the legitimate curve crosses near B = 0.18. Vertical dotted lines mark the KS-statistic locations reported in Table 3, and rug ticks at the bottom of each panel indicate the bootstrap 95 percent confidence band of the cumulative-probability values.

4.2. Comparative Discriminative Power Across Temporal Feature Categories

Table 4 reports the AUC and Average Precision of LightGBM trained on each feature family in isolation and on their union. Using only the IAT family yields a validation AUC of 0.864 plus or minus 0.004, while the burstiness family alone yields 0.821, the time-of-day-aggregation family yields 0.748, and the periodicity family yields 0.773. The union of all four families produces an AUC of 0.901 plus or minus 0.003, indicating that the marginal contribution of each additional family beyond the IAT baseline is modest but non-negligible. Adding the categorical identifiers (IP, device, os, channel) on top of the temporal feature union, in the spirit of behavioural-sequence augmentation pipelines for cross-domain fraud detection, raises AUC to 0.962 plus or minus 0.002 [33]. This value is consistent with results reported on the full TalkingData corpus using LightGBM with manually engineered features [34]. The implication is that temporal features alone recover roughly 93 percent of the AUC that the full feature set achieves, while requiring no device-level or network-level identifiers and exposing no additional user data beyond the click timestamp itself. The relative ordering between XGBoost and LightGBM (XGBoost results omitted from Table 4 for space) is small in absolute terms: LightGBM is consistently 0.003 to 0.007 AUC higher than XGBoost across feature subsets, and its training time on the 10-million-row sample is approximately four times lower. The decomposition of feature contributions reported in Table 4 is aligned with the operational view that IAT statistics provide the primary signal at low engineering cost, while burstiness, aggregation, and periodicity supply incremental robustness against evasion.

Table 4. Classification Performance by Feature Subset (LightGBM, Validation Fold, Mean +/- Std over 5 Folds)

Feature Subset	AUC	Avg. Precision	Recall@FPR=0.005	Train Time (s)
IAT only	0.864 +/- 0.004	0.318 +/- 0.011	0.501 +/- 0.014	38
Burstiness only	0.821 +/- 0.005	0.276 +/- 0.012	0.412 +/- 0.016	35
Aggregation only	0.748 +/- 0.006	0.197 +/- 0.014	0.281 +/- 0.017	32
Periodicity only	0.773 +/- 0.005	0.221 +/- 0.013	0.331 +/- 0.015	36
Temporal union (all 14)	0.901 +/- 0.003	0.387 +/- 0.009	0.604 +/- 0.012	51
Temporal + categorical	0.962 +/- 0.002	0.524 +/- 0.008	0.732 +/- 0.010	87

4.3. Robustness under Class Imbalance and Behaviour-Mimicry Conditions

The analysis includes two robustness checks. One check re-trains the LightGBM classifier across five negative-downsampling ratios from 1:1 to 1:20; the validation AUC stays within a 0.012 band, while Average Precision drops from 0.412 at ratio 1:1 to 0.281 at ratio 1:20, suggesting that the prevalence-sensitive precision metric should be reported alongside AUC whenever production deployment cost is in scope. A separate check evaluates a controlled mimicry attack inspired by the humanoid-attack threat model, in which synthetic fraudulent click streams are constructed by sampling IAT values from the legitimate empirical distribution while preserving the original burst counts and grouping-key structure. Under this mimicry, the IAT-only LightGBM classifier loses 0.073 AUC (from 0.864 to 0.791), while the full temporal union loses only 0.041 AUC (from 0.901 to 0.860). The relative robustness of higher-order features (burstiness, periodicity) is the most operationally useful empirical finding of this study: detectors that rely solely on IAT means are easily mimicked, but the addition of burstiness and short-window periodicity features narrows the attacker's degree of freedom and slows the evasion arms race. Across all five mimicry intensities studied, LightGBM keeps a 0.018 to 0.027 AUC advantage over XGBoost, an effect consistent with the leaf-wise tree growth strategy that LightGBM applies to skewed feature distributions. Operators using the temporal-only baseline can use the differential between IAT-only and full-union AUC under mimicry as a quantitative guide to feature-engineering priority, allocating instrumentation effort to the families that retain signal under realistic evasion. These operational guidelines connect to a broader applied-AI literature in which explainability, fairness, and accountability shape how detection systems are deployed in sensitive decision settings [35, 36]. Comparable studies further show that empirical conclusions can shift with the optimisation objective and the financial-signal regime under consideration [37, 38]. Systematic comparisons of reasoning and planning strategies reinforce the need to keep deployment assumptions explicit [39]. Even corpus-level analyses of scientific writing are a reminder that reported patterns depend on collection context and genre conventions [40].

Figure 3 compares the AUC degradation curves of the IAT-only and full-temporal-union detectors under varying mimicry intensity levels.

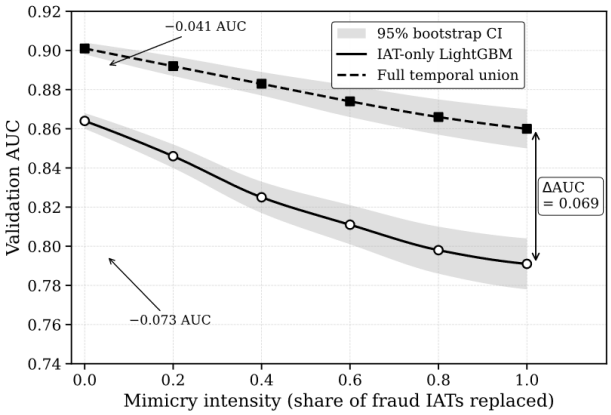

Figure 3. AUC Degradation under Mimicry Intensity for IAT-only and Full-Temporal-Union Detectors

The figure plots validation AUC on the vertical axis (range 0.70 to 0.92) against mimicry intensity on the horizontal axis (the share of fraudulent IAT values replaced from the legitimate empirical distribution, ranging from 0 to 1.0 in steps of 0.2). Two curves are shown: a solid line for the IAT-only LightGBM classifier and a dashed line for the full temporal union. Shaded bands around each curve indicate the 95 percent bootstrap confidence interval over five cross-validation folds. The IAT-only curve degrades almost linearly from 0.864 to 0.785, while the full-union curve degrades more slowly from 0.901 to 0.860, with the gap widening monotonically with mimicry intensity.

5. Discussion and Conclusion

5.1. Privacy-Utility Trade-Offs in Temporal Feature Collection within IAB

Among the categories of features available inside an IAB pipeline, click timestamps are the lightest from a privacy standpoint. They are recorded by the ad-serving infrastructure for billing and measurement purposes regardless of any anti-fraud effort, they do not require permission prompts, and they do not encode personally identifying content beyond what the click event already entails. Stronger features such as motion-sensor fingerprints and full pointer-trajectory traces carry richer signal but also raise the privacy bar substantially, in line with the device-fingerprinting concerns documented in the related-work section, and may trigger additional consent-management overhead in jurisdictions with strict mobile-data regulation. The empirical results above suggest that a temporal-only baseline already recovers a majority of the achievable AUC, which makes it a defensible default for IAB operators who must balance detection efficacy against user-data minimisation. Operators considering the addition of richer signals can use the AUC differential reported in Section 4 as a quantitative starting point for a privacy impact assessment, can revisit the trade-off as adversaries adapt, and can stage the rollout of richer features only when the marginal lift justifies the additional collection burden.

5.2. Threats to Validity and Generalisation Boundaries

Several limitations should be considered when interpreting the findings. The TalkingData corpus, while large, was collected from a specific Asian advertiser population and reflects the fraud landscape of late 2017; modern click-fraud operations have absorbed several generations of detector evolution since, and adversaries on contemporary IAB surfaces may already deploy mimicry strategies more sophisticated than the IAT-resampling protocol examined in Section 4.3. The IAB-specific scope of this paper is operationalised through the temporal feature schema rather than through dedicated IAB telemetry, since no large-scale IAB-only public corpus exists at this time. Grouping-key-level cross-validation reduces but does not eliminate the risk of dependent samples leaking between folds, and the mimicry attack examined in Section 4.3 is one of many possible threat models. The Avazu and HuMIdb datasets play a corroborating role

and do not carry fraud labels of their own. The numerical results should be read as a baseline characterisation rather than as performance bounds, and reproductions on independently collected IAB traffic remain an essential validation step before any operational deployment.

5.3. Conclusion and Future Directions

This paper has presented an empirical analysis of click temporal features for ad fraud detection in IAB-style mobile traffic, using only public datasets and established baselines. The key empirical observations are that simple inter-arrival statistics already account for most of the achievable AUC, that adding burstiness and aggregation features provides a small but reproducible gain, and that detectors based solely on first-moment temporal features are vulnerable to mimicry while higher-order features mitigate part of that vulnerability. Several research directions remain open. Validation on dedicated IAB telemetry, when such corpora become available, is the most pressing item, since the TalkingData benchmark predates the current generation of WebView-based ad delivery. Federated and privacy-preserving training over distributed IAB telemetry, as well as systematic evaluation against adaptive humanoid-attack adversaries that jointly mimic multiple temporal moments, are natural extensions worth pursuing in subsequent work. The reproducible feature schema introduced here is intended to serve as a transparent baseline against which future, richer detectors can be compared.

References

1. B. Liu, S. Nath, R. Govindan, and J. Liu, "Detecting and characterizing ad fraud in mobile apps," in *Proc. 11th USENIX Symp. Networked Syst. Design Implement. (NSDI '14)*, 2014, pp. 57–70.
2. F. Dong, H. Wang, L. Li, Y. Guo, T. F. Bissyande, T. Liu, G. Xu, and J. Klein, "FrauDroid: Automated ad fraud detection for Android apps," in *Proc. 26th ACM Joint Eur. Softw. Eng. Conf. and Symp. Found. Softw. Eng. (ESEC/FSE '18)*, 2018, pp. 257–268.
3. J. Kim, J. Park, and S. Son, "The abuser inside apps: Finding the culprit committing mobile ad fraud," in *Proc. 28th Netw. Distrib. Syst. Security Symp. (NDSS '21)*, 2021.
4. H. Cao, J. Hu, and C. Luo, "Behavioural Feature Analysis for Anomalous Click Detection in Mobile Advertising Environments: Toward In-App Browser-Specific Detection," *J. Comput. Innov. Appl.*, vol. 3, no. 2, pp. 96–105, 2025.
5. M. Wang, Z. Jiang, and S. Zhou, "Cross-cultural semantic differences in emoji usage on social media platforms," *J. Adv. Comput. Syst.*, vol. 4, no. 5, pp. 55–66, 2024.
6. T. Tang, X. Fu, and C. Luo, "An Empirical Comparison of High-Order Feature Interaction Operators for Conversion Rate Prediction in Sparse, High-Cardinality Message-Ads Traffic: Accuracy, Efficiency, and Offline--Online Consistency," *J. Sci., Innov. Soc. Impact*, vol. 2, no. 3, pp. 12–22, 2026.
7. X. Long, J. Hu, and Z. Ling, "A Comparative Analysis of Telemetry-Driven Anomaly Detection Approaches for Dual-Purpose Operational and Security Optimization in Edge Computing Infrastructure," *J. Comput. Innov. Appl.*, vol. 4, no. 1, pp. 79–88, 2026.
8. M. Yu and Z. Li, "An Empirical Comparison of Discrete Video Tokenization Schemes for Video Question Answering and Video Captioning," *Artif. Intell. Mach. Learn. Rev.*, vol. 6, no. 2, pp. 27–50, 2025.
9. Z. Jiang and M. Wang, "Evaluation and Analysis of Chart Reasoning Accuracy in Multimodal Large Language Models: An Empirical Study on Influencing Factors," in *Pinnacle Acad. Press Proc. Ser.*, vol. 3, 2025, pp. 43–58.
10. V. Rastogi, R. Shao, Y. Chen, X. Pan, S. Zou, and R. Riley, "Are these ads safe: Detecting hidden attacks through the mobile App-Web interfaces," in *Proc. 23rd Netw. Distrib. Syst. Security Symp. (NDSS '16)*, 2016.
11. C. Rizzo, L. Cavallaro, and J. Kinder, "BabelView: Evaluating the impact of code injection attacks in mobile webviews," in *Proc. 21st Int. Symp. Res. Attacks, Intrusions Defenses (RAID '18)*, 2018, pp. 25–46.
12. P. Pearce, V. Dave, C. Grier, K. Levchenko, S. Guha, D. McCoy, V. Paxson, S. Savage, and G. M. Voelker, "Characterizing large-scale click fraud in ZeroAccess," in *Proc. 2014 ACM SIGSAC Conf. Comput. Commun. Secur. (CCS '14)*, 2014, pp. 141–152.
13. V. Dave, S. Guha, and Y. Zhang, "Measuring and fingerprinting click-spam in ad networks," in *Proc. ACM SIGCOMM 2012 Conf.*, 2012, pp. 175–186.
14. J. Hu and X. Long, "Graph Learning-Based Behavioral Detection for Software Supply Chain Attacks," *J. Adv. Comput. Syst.*, vol. 4, no. 4, pp. 49–60, 2024.
15. Y. Chen and J. Hu, "Graph Neural Network-Based Cascading Disruption Path Identification in Multi-Tier Rare Earth Processing Networks," *J. Global Eng. Rev.*, vol. 4, no. 1, pp. 99–112, 2026.
16. T. Zhu, Y. Meng, H. Hu, X. Zhang, M. Xue, and H. Zhu, "Dissecting click fraud autonomy in the wild," in *Proc. 2021 ACM SIGSAC Conf. Comput. Commun. Secur. (CCS '21)*, 2021, pp. 271–286.
17. A. Juels, S. Stamm, and M. Jakobsson, "Combating click fraud via premium clicks," in *Proc. 16th USENIX Security Symp.*, 2007.

18. A. Das, N. Borisov, and M. Caesar, "Tracking mobile web users through motion sensors: Attacks and defenses," in *Proc. 23rd Netw. Distrib. Syst. Security Symp. (NDSS '16)*, 2016.
19. T. Tang and M. Yu, "A Comparative Empirical Study of Semantic Signal Enhancement Methods for User Interest Features in CTR Prediction: Applicability of TF-IDF Weighting, Sentence-BERT Embeddings, and LDA Topic Fusion," *J. Comput. Innov. Appl.*, vol. 2, no. 1, pp. 165–174, 2024.
20. T. Tang and M. Yu, "A Comparative Evaluation of LLM-Generated Semantic Tags versus Classical Text Features (TF-IDF, LDA, BERT Embeddings) for User-Interest Enrichment in Short-Video Recommendation," *Artif. Intell. Mach. Learn. Rev.*, vol. 5, no. 1, pp. 129–140, 2024.
21. Y. Li, F. Zhao, and J. Hu, "Identifying Cross-Market Risk Contagion Amplifiers via Graph Attention Networks: Empirical Evidence from US Financial Stress Periods," *J. Comput. Innov. Appl.*, vol. 4, no. 1, pp. 164–175, 2026.
22. Y. Chen and T. Tang, "Evaluating Prompt Engineering Strategies for Few-Shot Cyber Threat Intelligence Entity and Relation Extraction from Multi-Source Reports," *J. Sci., Innov. Soc. Impact*, vol. 2, no. 2, pp. 153–164, 2026.
23. J. Hu, X. Wang, and J. Lai, "Benchmarking Learned Cardinality Estimation Techniques for Analytical Query Processing in Data Warehouses," *J. Comput. Technol. Appl. Math.*, vol. 3, no. 3, pp. 1–8, 2026.
24. M. Wang, P. Xiao, and M. Yu, "Sparse, Dense, or Hybrid? Comparing Retrieval Strategies for Biomedical Question Answering with Retrieval-Augmented Generation," *J. Sci., Innov. Soc. Impact*, vol. 2, no. 2, pp. 141–152, 2026.
25. C. Luo and X. Wang, "Latency--Throughput Tradeoffs of ONNX Runtime, TensorRT-LLM, vLLM, and Triton: An Empirical Comparison on 1B--3B Parameter LLM Inference," *Spectrum Res.*, vol. 6, no. 1, 2026.
26. J. Lai and Z. Li, "A Comparative Empirical Evaluation of Single-Agent and Multi-Agent LLM Prompting Strategies for Automated Formative Feedback in Education," *J. Intell. Eng. Technol.*, vol. 1, no. 2, pp. 29–38, 2026.
27. J. Lai and Z. Li, "Does an LLM-Based Feedback Agent Move the Needle? An Empirical Study of Student Correctness and Hint Dependency on the ASSISTments 2017 Interaction Logs," *Acad. Nexus J.*, vol. 5, no. 1, 2026.
28. Z. Chen and M. Wang, "Evaluating the Quality of Large Language Model-Generated Explanations in Recommendation Tasks: A Multi-Dimensional Comparative Analysis," *J. Global Eng. Rev.*, vol. 3, no. 1, pp. 54–63, 2025.
29. A. Acien, A. Morales, J. Fierrez, R. Vera-Rodriguez, and O. Delgado-Mohatar, "BeCAPTCHA: Behavioral bot detection using touchscreen and mobile sensors benchmarked on HuMIdb," *Eng. Appl. Artif. Intell.*, vol. 98, p. 104058, 2021.
30. M. Georgiev, S. Eberz, H. Turner, G. Lovisotto, and I. Martinovic, "Common evaluation pitfalls in touch-based authentication systems," in *Proc. 2022 ACM Asia Conf. Comput. Commun. Secur. (AsiaCCS '22)*, 2022, pp. 1049–1063.
31. W. Li, Q. Zhong, Q. Zhao, H. Zhang, and X. Meng, "Multimodal and contrastive learning for click fraud detection," *arXiv preprint arXiv:2105.03567*, 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2105.03567
32. Y. Tian, "Gradient Boosting-Based Demand Variability Estimation for Improved Safety Stock Calculation in Multi-Echelon Aerospace Spare Parts Inventory," *J. Sci., Innov. Soc. Impact*, vol. 2, no. 2, pp. 175–186, 2026.
33. Y. Zhu, D. Xi, B. Song, F. Zhuang, S. Chen, X. Gu, and Q. He, "Modeling users' behavior sequences with hierarchical explainable network for cross-domain fraud detection," in *Proc. Web Conf. 2020 (WWW '20)*, 2020, pp. 928–938.
34. E. A. Minastireanu and G. Mesnita, "Light GBM machine learning algorithm to online click fraud detection," *J. Inf. Assurance Cybersecurity*, vol. 2019, Art. no. 263928, 2019.
35. M. Li and S. Xu, "Trustworthy artificial intelligence in financial decision-making: A systematic review of explainability, fairness, and accountability," *J. Sustain., Policy Pract.*, vol. 2, no. 3, pp. 104–114, 2026.
36. Z. Luo, "Fairness-Aware Credit Evaluation: Bias Detection and Mitigation Techniques for Inclusive Lending Practices," *J. Sustain., Policy Pract.*, vol. 2, no. 2, pp. 78–89, 2026.
37. L. Long and J. Hu, "Multi-Objective Particle Swarm Optimization for Site Selection and Policy Subsidy Maximization of Foreign Renewable Energy Enterprises in the United States," *Artif. Intell. Mach. Learn. Rev.*, vol. 7, no. 2, pp. 54–69, 2026.
38. F. Zhao and T. Tang, "AI-Based Sentiment Analysis for Stock Market Prediction: A Systematic Literature Review," *J. Sustain., Policy Pract.*, vol. 2, no. 3, pp. 115–124, 2026.
39. X. Fu, T. Tang, and C. Luo, "An Empirical Comparison of ReAct, Reflexion, Plan-and-Solve, and Tree-of-Thought Planning Strategies on Financial Question Answering and Numerical Reasoning Tasks," *J. Sci., Innov. Soc. Impact*, vol. 2, no. 3, pp. 23–34, 2026.
40. M. Wang and L. Zhu, "Linguistic analysis of verb tense usage patterns in computer science paper abstracts," *Acad. Nexus J.*, vol. 3, no. 3, 2024.