Lightweight SPIT Detection Under Adversarial Mimicry

Antonio Jorge Andrade[1], Luciano Barbosa[1]

CESAR School, Brazil
ajbloa@cesar.school, lab2@cesar.org.br

Abstract. Spam over Internet Telephony (SPIT) detection faces escalating adversarial evasion via behavioral mimicry, while deep learning defenses often struggle to meet real-time scalability and transparency requirements. We formalize the "Fraudster's Trilemma" and evaluate lightweight classifiers over an eight-dimensional feature space spanning physical, volumetric, and behavioral planes, using a simulation of 500,000 subscriber profiles across three stages of adversarial complexity (E_1–E_3). Ablation evidence indicates that the behavioral vector collapses under calibrated mimicry ($F_{0.5} = 0.03$ under Random Forest at E_3), while the infrastructure-bound physical vector sustains an independent boundary of $F_{0.5} = 0.98$; the volumetric vector collapses under a linear boundary but persists under tree ensembles, which recover the class boundary from dispersion statistics that this benchmark does not equalize. Lightweight ensembles sustain high precision on this benchmark while scoring each subscriber independently, and per-decision SHAP attributions give operators an auditable justification for each blocking decision.

CCS Concepts: • **Computing methodologies** → **Machine learning algorithms**.

Keywords: adversarial mimicry, anomaly detection, machine learning, SPIT detection

1. INTRODUCTION

Telecommunications fraud is undergoing a structural transition. Caller identity authentication frameworks such as STIR/SHAKEN cryptographically attest the originating identity of SIP calls, progressively narrowing the Calling Line Identification (CLI) spoofing vector. This reshapes the adversary's cost structure: where identity can no longer be forged in signaling, it must be acquired legitimately. A rational response is *Local Presence Dialing*—operating distributed fleets of physical SIM cards inside the target network, so that malicious traffic originates from cryptographically authenticated, legitimate subscriber profiles. This induces a critical detection asymmetry: signaling-layer defenses lose purchase by construction, forcing detection onto the behavioral and physical analysis of Call Detail Records (CDRs).

Simultaneously, regulatory frameworks such as the GDPR in Europe and the LGPD in Brazil subject automated decisions that block a subscriber's service to transparency and review obligations. While complex deep learning architectures—particularly Graph Neural Networks (GNNs)—are frequently proposed for SPIT detection due to their high performance, they introduce two substantial barriers for regulated deployment. First, GNN deployment carries a structural cost that feature-based scoring does not: a GNN relates subscribers to one another rather than scoring them in isolation, and inline operation requires maintaining that relational structure as a live, stateful artifact. The classifiers evaluated here score each subscriber profile independently from a fixed-width feature vector, incurring no shared state between decisions. Second, GNN decision processes are typically non-transparent, rarely producing auditable, per-decision justifications. LGPD Article 20 grants data subjects the right to request review of decisions taken solely on automated processing that affect their interests, and GDPR Article 22 constrains such decisions while requiring meaningful information about the logic involved. Whether these provisions amount to a right to explanation of individual decisions remains contested; regardless of that reading, an operator that cannot articulate why a specific subscriber was blocked is poorly positioned to satisfy a review request. Both barriers are architectural and regulatory rather than empirical: we do not claim superiority over trained GNN baselines, which fall outside the scope of this evaluation.

To evaluate the physical limits of adversarial evasion, we formalize these boundaries through the *Fraudster's Trilemma*. This framework posits that an adversary operating a Local Presence SPIT campaign must choose among three unfavourable options: (a) remaining physically immobile, exposing the infrastructure to geographic tracking; (b) physically moving the infrastructure to mimic human mobility, incurring substantial operational overhead; or (c) reducing call volumes to human scales, which substantially undermines the campaign's financial viability. Based on this trilemma, we formulate three core hypotheses: **H1 (Software Mimicry)**, stating that behavioral and volumetric vectors can be neutralized via software adjustments; **H2 (Hardware Asymmetry)**, stating that the physical mobility vector remains robust under extreme mimicry; and **H3 (Algorithmic Sufficiency)**, stating that lightweight classifiers are computationally sufficient, without the need for complex architectures.

Proposed Solution and Contributions. We propose an explainable multi-vector SPIT detection framework built on infrastructure hardware constraints, whose eight-dimensional feature space spans the physical, volumetric, and behavioral planes and is evaluated on a synthetic benchmark of 500,000 subscriber profiles. The contributions are threefold: (1) we provide ablation evidence indicating that the behavioral vector collapses under calibrated mimicry while the physical vector remains robust and grows monotonically in attributed importance; (2) we report evidence for the algorithmic sufficiency of lightweight ensemble classifiers on this benchmark, sustaining high precision boundaries without maintaining a relational structure at inference time; and (3) we show that the resulting decisions can be accompanied by per-decision SHAP attributions, giving operators a concrete artifact with which to respond to review and transparency requests.

2. RELATED WORK

State-of-the-art SPIT detection techniques span three technical families, as catalogued in a survey of the area [Azad et al. 2018]. Rule-based systems apply heuristic thresholds on metrics such as call rates or session durations; while computationally trivial, they degrade significantly when confronted with calibrated human mimicry. Deep learning methods, particularly GNNs and autoencoders, achieve competitive macro-performance in static benchmarks [Hu et al. 2023] but face multi-million-subscriber graph-construction bottlenecks [Zhong et al. 2024] and sit uneasily with the transparency obligations of GDPR Article 22 and LGPD Article 20. A third family operates directly on CDR metadata at carrier scale: Liu et al. mine call detail records to expand telephone spam blacklists, tuning the detector for zero false positives on the explicit grounds that carriers cannot absorb the cost of blocking a legitimate subscriber [Liu et al. 2018]. That asymmetric cost structure is the same one that motivates our choice of $F_{0.5}$ as the primary metric. Inspired by the scalability of classical tabular classifiers, this paper proposes an alternative middle ground, leveraging multi-vector feature extraction executed with lightweight ensemble algorithms [Breiman 2001; Ke et al. 2017].

Closest to the physical plane evaluated here is work that keys on infrastructure rather than on behavior. Reaves et al. block cellular interconnect bypass at the network edge, a setting in which the adversary likewise operates a fixed installation of physical SIMs inside the target network [Reaves et al. 2015]. The threat model is adjacent rather than identical—bypass fraud monetizes call termination, not the delivery of a solicitation—but it shares the property that this paper tests under mimicry: hardware placement is expensive for an adversary to falsify.

Table I summarizes how this paper positions itself against the closest prior work. While contemporary literature has explored isolated volumetric or behavioral planes against naive attackers, to the best of our knowledge, no prior study evaluates the resilience of an infrastructure-bound physical vector against an adversary executing progressive, humanized behavioral and volumetric mimicry, while simultaneously preserving full decision-level auditability.

A structural gap shared across these methodologies is that experimental benchmarks typically as-

Table I. Positioning against closest prior work in SPIT detection. A dash denotes a dimension not addressed by the cited work.

Work	Interpretable	Real-Time Deployable	Multi-Vector Evaluation	Adversarial Mimicry
Rule-based Systems	Yes	Yes	Partial	–
[Hu et al. 2023]	Partial	Partial	Yes	–
[Liu et al. 2018]	–	–	Partial	–
[Reaves et al. 2015]	–	–	Partial	–
This work	**Yes**	**Yes**	**Yes**	**Yes**

sume a naive adversary making no systematic attempt to mask traffic signatures. This paper targets that gap directly; our results indicate that physical network indicators remain largely invariant even as software-driven topological counters degrade substantially in attributed importance.

3. METHODOLOGY

The proposed detection methodology is structured sequentially across two phases, shown in Figure 1. Phase 1 (Data) orchestrates the synthetic generation of network interactions and their structural formatting into CDRs; Phase 2 (Machine Learning) executes multi-vector feature engineering across the physical, volumetric, and behavioral planes and, once these feature vectors are consolidated, performs model training, validation, and the evaluation of lightweight classifiers (Random Forest, LightGBM, and Logistic Regression).

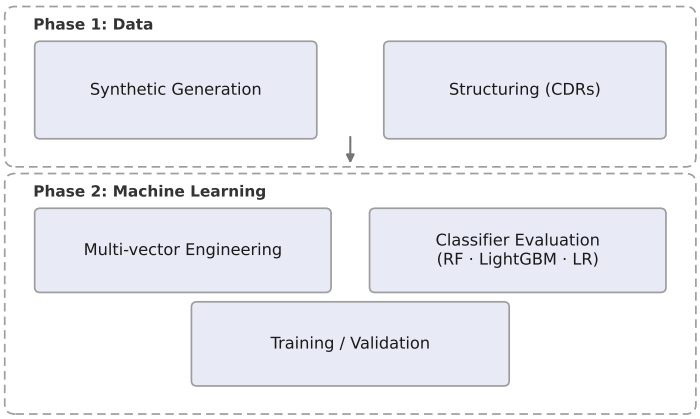

Fig. 1. Overview of the proposed SPIT detection pipeline.

3.1 Adversary Space

We model the adversary along two orthogonal axes: *Volumetric* (the rate and timing of calls) and *Behavioral* (the topological distribution of destinations). Table II maps three progressive, cumulative stages of mimicry complexity (E_1, E_2, E_3): the behavioral vector is held at high fidelity throughout via Zipf-calibrated destinations, while the volumetric vector is driven from a naive attack profile down to a human-indistinguishable one. Holding the software-level axes at their strongest setting isolates the discriminatory boundary of the infrastructure-bound constraints.

Our threat model focuses on the Local Presence SPIT adversary, which operates physical SIM cards inside the target network. The adversary employs a stationary *simbox* (SIM bank) located at a fixed geographic coordinate. We stress that adversary immobility is not a simplifying assumption but the operational core of the Fraudster's Trilemma: large-scale simbox deployments require stable power, antenna arrays, and backhaul connectivity, rendering sustained physical relocation costly at scale. Operator-scale CDR evidence corroborates this constraint in the adjacent setting of interconnect

bypass fraud, where fraudulent simboxes exhibit near-static physical locations and disproportionately high outgoing-call ratios [Murynets et al. 2014]. The simbox connects to the nearest available cellular base station (ERB).

3.2 Synthetic Data Generation

We generate a synthetic dataset of 500,000 subscriber profiles (495,000 legitimate and 5,000 SPIT) spanning a seven-day observation window. The reliance on synthetic data is a methodological necessity rather than a convenience: publicly available CDR datasets typically strip geolocation telemetry and ERB identifiers to comply with privacy and secrecy mandates (LGPD/GDPR), precluding evaluation of the physical plane. We therefore ground the simulation in parametric distributions and kinematic constraints established in the human-mobility literature [Calabrese et al. 2013; Forghani et al. 2020]. Legitimate profiles are generated using a multi-modal agent-based discrete event simulator, which models distinct demographic segments (e.g., business users, personal users, call centers). SPIT profiles are generated by applying the three-stage adversary model (E_1, E_2, E_3) to a base simbox profile. To improve behavioral realism and statistical fidelity without relying on generic randomized heuristics, the agents' activities are governed by explicit parametric probability distributions:

—**Volumetric and Temporal Modeling:** The call frequency V per subscriber is drawn from a Log-Normal distribution, $V \sim \min(\max(\lfloor \exp(\mu + \sigma \cdot Z) \rfloor, V_{\min}), V_{\max})$, where $Z \sim \mathcal{N}(0,1)$. The parameters $(\mu, \sigma, V_{\min}, V_{\max})$ shift across experimental stages (E_1–E_3) to simulate progressive mimicry. For human profiles, $\mu = 3.0$ and $\sigma = 1.0$. Temporally, Inter-Arrival Times (IAT, Δt) follow a multi-modal mixture of uniform continuous distributions, $\Delta t \sim \alpha \cdot \mathcal{U}(t_{\text{win\_min}}, t_{\text{win\_max}}) + (1-\alpha) \cdot \mathcal{U}(t_{\text{off\_min}}, t_{\text{off\_max}})$, where $\alpha = 0.75$ isolates active communication windows ($30 \leq t_{\text{win}} \leq 3600$ s) from nocturnal idle periods.

—**Topological Destinations:** To replicate human social network structures, destination selection for directory size $K \sim \mathcal{U}(5, 20)$ is bound to Zipf's Law: $P(k) = (1/k^a)/\sum_{m=1}^{K}(1/m^a)$. The scaling exponent is fixed at $a = 1.5$ for humans, while SPIT adversaries manipulate this topology dynamically by sampling $a_{\text{spit}} \sim 2.0 \cdot \text{Beta}(2,3) + 0.5$ to introduce subtle variations into destination entropy metrics.

—**Spatial Mobility:** Legitimate agents transition between localized spatial clusters (Home, Work, Anchors) via a localized Bernoulli handover process, $h \sim \text{Bernoulli}(p_{\text{handover}})$. Conversely, the SPIT adversary operates from a geographically stationary simbox profile, attempting to evade detection by alternating connections among adjacent sectors ($0.20 \leq p_{\text{swap}} \leq 0.45$) or routing traffic via regional VoIP nodes. For all agents, isotropic Gaussian noise ($\sigma_{\text{geo}} \approx 10$ m) is injected into the coordinates to represent residual positioning error, a known confound when reconstructing trajectories from cellular data [Forghani et al. 2020].

3.3 Feature Engineering

The feature engineering pipeline maps subscriber behavior into an eight-dimensional interpretable metric space. Interpretability here is a property of the feature space itself rather than of any post-hoc attribution method: every feature is a named physical or statistical quantity that an operator can inspect directly, and no uninterpretable high-dimensional projection is used. Two further features

Table II. Taxonomy of the evaluated adversary space, illustrating the cumulative progression of evasive mimicry across volumetric and behavioral dimensions (Stages E_1 to E_3).

Stage	Volumetric Mimicry	Behavioral Mimicry
E_1	None (naive volumetric attack)	High (Zipf-calibrated destinations)
E_2	Partial (Intermediate volume)	High (Zipf-calibrated destinations)
E_3	High (Human-indistinguishable volume)	High (Zipf-calibrated destinations)

produced by the extraction pipeline were withheld from every experiment reported here: both separate the classes almost perfectly, but they do so by exposing the remote-VoIP routing signature rather than the Local Presence behavior under study. The remaining features are partitioned into the three semantic planes of Table III, which isolate behavioral fraud-evasion layers from infrastructure constraints.

The physical plane carries two implementation decisions that the table does not convey. Displacement between sequential ERBs is computed with an equirectangular approximation rather than a geodesic one: it requires no iterative solution, reducing per-profile extraction cost, and its departure from a geodesic distance is negligible at the displacement scale the detector keys on. The Spatial Impossibility Ratio thresholds apparent speed at 1,200 km/h—a deliberately permissive bound, set above any conventional means of transport so that the feature fires only on displacements no physical subscriber could plausibly have performed—and transitions separated by less than one second or less than 200 m are held at zero speed, so that positioning jitter cannot manufacture spurious velocities.

Table III. Multi-Vector Feature Engineering Pipeline. The physical plane is not falsifiable in software; the volumetric plane is falsifiable at the cost of throughput, and the behavioral plane at the cost of campaign reach.

Plane	Formalized Feature	Technical Description & Intent
Physical	Median Apparent Speed	Displacement velocity between successive cell tower handovers.
	Maximum Apparent Speed	Peak structural velocity, isolating single anomalous relocations.
	Spatial Impossibility Ratio	Fraction of cell-site transitions violating human spatiotemporal limits.
Volumetric	Total Call Count	Network transaction frequency over the observation window.
	Call Duration Dispersion	Standard deviation of call lengths, capturing automated cut-offs.
	IAT Standard Deviation	Dispersion of call intervals, isolating rhythmic dialing scripts.
Behavioral	Unique Destination Count	Cardinality of the subscriber's calling network.
	Destination Shannon Entropy	Structural uncertainty and concentration of destination routing.

4. EXPERIMENTAL SETUP

Data partitioning and training configuration. The synthetic datasets generated across the three cumulative adversarial stages (E_1, E_2, and E_3) were independently partitioned by stratified random splitting: 20% of instances were held out for final evaluation, 60% were used for training, and the remaining 20% were reserved as a validation partition not consumed by the results reported here. Stratification preserves the class imbalance ratio (approximately 99:1) across all partitions. Model stability was additionally checked by five-fold stratified cross-validation on the development partition, which returned an ROC AUC above 0.999 with a standard deviation below 0.001 across folds at every stage. Profiles with fewer than ten call transactions were excluded prior to training, since dispersion and entropy statistics are unstable at that support; the filter removes roughly one quarter of legitimate profiles at every stage and a comparable share of E_3 SPIT profiles, whose volumes match the human distribution by construction, leaving approximately 380,000 profiles evaluated per stage. A fixed seed (`random_state = 42`) was used to support reproducibility of the reported runs.[1]

Classifier configuration. Three classifiers spanning complementary algorithmic families were evaluated: Logistic Regression (linear analytical baseline), Random Forest (bagging ensemble, primary ar-

[1]The synthetic data generator, feature-extraction pipeline, classifier configurations, and evaluation scripts are publicly available at `https://github.com/AntonioJorg/SPITbyCDR`.

Table IV. Classification performance across progressive adversarial stages (E_1 to E_3).

Stage	Classifier Architecture	Precision	Recall	$F_{0.5}$
E_1	Logistic Regression	0.94	1.00	0.95
	LightGBM	1.00	1.00	1.00
	Random Forest	**1.00**	**1.00**	**1.00**
E_2	Logistic Regression	0.93	0.99	0.94
	LightGBM	1.00	1.00	1.00
	Random Forest	**1.00**	**1.00**	**1.00**
E_3	Logistic Regression	0.97	0.99	0.97
	LightGBM	1.00	0.99	1.00
	Random Forest	**1.00**	**0.99**	**1.00**

chitecture), and LightGBM (gradient-boosting ensemble). All three score a profile from a fixed-width feature vector in a single scoring pass, with no dependency on other subscribers' records at inference time. All models were configured with automatic class balancing (`class_weight="balanced"`; `"balanced_subsample"` for Random Forest, applied per bootstrap draw), which weights classes inversely to their frequency so that the minority class is not dominated during fitting.

Evaluation metrics. Performance is evaluated using the $F_{0.5}$ score, formalized as $F_\beta = (1 + \beta^2) \cdot (\text{Precision} \cdot \text{Recall})/[(\beta^2 \cdot \text{Precision}) + \text{Recall}]$ with $\beta = 0.5$. This metric prioritizes Precision over Recall to reflect the asymmetric cost matrix of telecommunications operators, where a false positive (blocking a legitimate subscriber) creates regulatory exposure (LGPD/GDPR) and churn risk.

5. RESULTS

Table IV reports classifier performance across the three stages of adversary evasion capability (E_1, E_2, E_3), measured on the holdout partition under a production-grade class imbalance (∼1% fraud prevalence).

Empirical evidence indicates that lightweight ensemble architectures sustain high precision boundaries (1.00) under extreme behavioral and volumetric mimicry (E_3). Under this configuration, both ensembles suppress false positives under advanced evasion. The saturation of these scores should be read against the synthetic origin of the benchmark: the generator fixes the physical separability that the ensembles exploit. No GNN baseline was trained on this benchmark, so no empirical comparison is claimed; the architectural contrast is the one set out in Section 1, and it holds here in that each profile is scored independently, with no relational state retained between decisions.

To contextualize these decision boundaries, Figure 2 tracks the mean absolute SHAP feature importance mapped across the three semantic detection planes.

The behavioral plane stays flat near zero (≈ 0.01) across all three stages. The volumetric plane decays steadily, from 0.33 at E_1 to 0.24 at E_3, as the attack profile aligns with human call volumes, while the infrastructure-bound physical plane grows monotonically over the same range, from 0.17 to 0.26. The two trajectories cross around E_2, where their attributed weights are nearly equal (physical 0.258, volumetric 0.255); the ordering after the crossing, not its exact location, is what the progression supports. SHAP values reflect model-specific attribution rather than causal effect, so we read this reversal as consistent with the software–hardware asymmetry rather than as a measurement of it: an adversary can adjust software-driven traffic rates at will, but remains bound to the physical placement of the cellular infrastructure it transmits through.

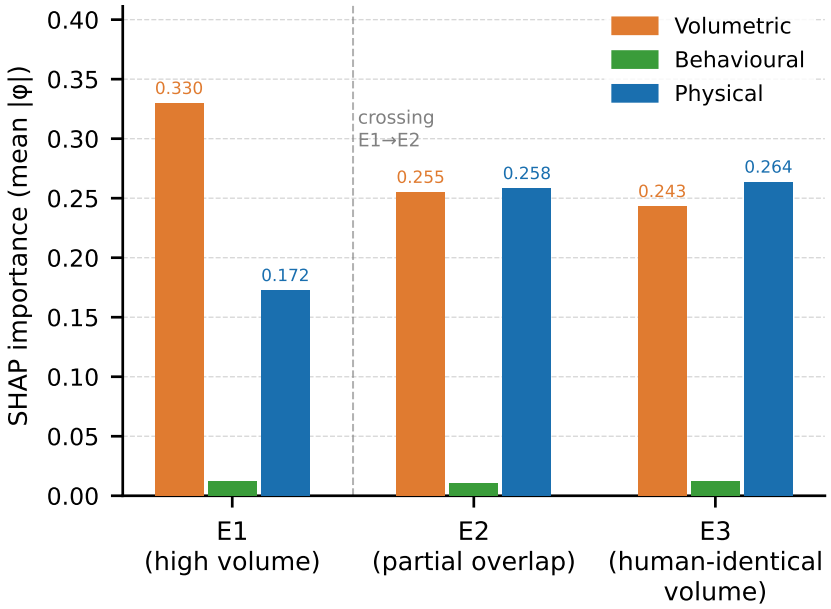

Fig. 2. Progression of mean absolute SHAP values (Random Forest), averaged over the features of each detection plane, across adversarial stages.

5.1 Vector Ablation Study

The vector ablation results summarized in Table V quantify this architectural asymmetry under extreme adversarial mimicry (Stage E_3).

Table V. Vector ablation analysis ($F_{0.5}$) under extreme adversarial mimicry (Stage E_3).

Isolated Semantic Plane	**Random Forest**	**LightGBM**	**Logistic Reg.**
Behavioral Plane Only	0.03	0.16	0.02
Volumetric Plane Only	0.99	0.93	0.20
Physical Plane Only	0.98	0.99	0.98
Behavioral + Volumetric	0.99	0.94	0.20
Physical + Volumetric	1.00	1.00	0.98
Full Feature Matrix (All Planes)	**1.00**	**1.00**	**0.97**

The behavioral plane degrades monotonically as mimicry advances. Isolated, it falls from $F_{0.5} = 0.33$ at E_1 to 0.13 at E_2 and 0.03 at E_3 under Random Forest: once destinations are drawn from a Zipf distribution calibrated to human social structure, destination entropy and unique-destination counts retain almost no separating signal. Adding the behavioral plane to the volumetric one leaves performance unchanged under Random Forest (0.99 against 0.99), indicating that the two are not complementary. The physical plane, by contrast, sustains an independent boundary of $F_{0.5} = 0.98$ across the same progression, consistent with H2.

The volumetric plane behaves differently across model families, and the difference is informative. Under Logistic Regression it collapses as H1 anticipates, falling from 0.80 at E_1 to 0.20 at E_3; under the tree ensembles it persists (0.99 and 0.93). The residual signal is therefore nonlinear. Stage E_3 equalizes the *number* of calls to the human distribution—which is what a linear boundary over this plane principally exploits—but not the dispersion of call durations or inter-arrival times, and the ensembles recover the class boundary from those dispersion statistics. H1 is consequently supported for the behavioral vector and for the linear treatment of the volumetric one; an adversary matching human

dispersion, and not merely human throughput, is not represented in this benchmark. Equalizing dispersion at E_3 is the immediate next step; the physical-plane result is unaffected, being measured with the volumetric features withheld. This also reconciles the ablation with Figure 2: the volumetric plane's declining SHAP share measures its contribution *within* the full model, where the physical plane progressively accounts for the same separation, and not its standalone discriminative power.

6. CONCLUSION

The results of this paper suggest that, for the Local Presence attack class evaluated here, complex deep learning architectures may be unnecessary—on both computational and regulatory grounds—for robust SPIT detection, although validation on operator-private CDRs remains necessary to confirm external validity. Under the Fraudster's Trilemma, while behavioral mimicry neutralizes destination entropy—the behavioral plane alone falls to $F_{0.5} \leq 0.16$ at E_3—the infrastructure-bound physical vector remains robust, scaling monotonically to the largest SHAP share of the three planes (0.26) in extreme scenarios (E_3). The volumetric plane degrades comparably under a linear classifier but not under tree ensembles, which recover the boundary from dispersion statistics that this benchmark's E_3 stage does not equalize. Lightweight ensemble architectures sustain high precision boundaries (1.00) under a production-grade class imbalance while scoring each subscriber independently, with no relational structure held as live state at inference time. Per-decision SHAP attributions give an operator a concrete artifact with which to answer a review request under LGPD Article 20 or an information obligation under GDPR Article 22; whether either provision compels an explanation of an individual decision remains contested, and we make no claim as to legal sufficiency. Future work will expand this framework toward vehicular-mobility adversaries and privacy-preserving cross-operator collaborative anomaly detection architectures.

REFERENCES

Azad, M. A., Morla, R., and Salah, K. Systems and methods for SPIT detection in VoIP: Survey and future directions. *Computers & Security* vol. 77, pp. 1–20, 2018.

Breiman, L. Random Forests. *Machine Learning* 45 (1): 5–32, 2001.

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira Jr., J., and Ratti, C. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies* vol. 26, pp. 301–313, 2013.

Forghani, M., Karimipour, F., and Claramunt, C. From cellular positioning data to trajectories: Steps towards a more accurate mobility exploration. *Transportation Research Part C: Emerging Technologies* vol. 117, pp. 102666, 2020.

Hu, X., Chen, H., Liu, S., Jiang, H., Wang, K., and Wang, Y. Who are the evil backstage manipulators: Boosting graph attention networks against deep fraudsters. *Computer Networks* vol. 232, pp. 109848, 2023.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 30. Curran Associates, Inc., Red Hook, NY, USA, 2017.

Liu, J., Rahbarinia, B., Perdisci, R., Du, H., and Su, L. Augmenting Telephone Spam Blacklists by Mining Large CDR Datasets. In *Proceedings of the 2018 ACM Asia Conference on Computer and Communications Security (ASIACCS)*. ACM, Incheon, Republic of Korea, pp. 273–284, 2018.

Murynets, I., Zabarankin, M., Piqueras Jover, R., and Panagia, A. Analysis and detection of SIMbox fraud in mobility networks. In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*. IEEE, Toronto, ON, Canada, pp. 1519–1526, 2014.

Reaves, B., Shernan, E., Bates, A., Carter, H., and Traynor, P. Boxed Out: Blocking Cellular Interconnect Bypass Fraud at the Network Edge. In *Proceedings of the 24th USENIX Security Symposium*. USENIX Association, Washington, D.C., USA, pp. 833–848, 2015.

Zhong, M., Lin, M., Zhang, C., and Xu, Z. A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges. *Computers & Security* vol. 141, pp. 103821, 2024.