



---

# Probabilistic Attribution Models for Digital Out-of-Home Advertising: A Design Science Approach to Bridging Physical Exposure and Digital Behavior

Muthupalaniappan Ramanathan

Submitted:03/02/2026

Revised: 08/03/2026

Accepted: 20/03/2026

**Abstract:** However, since no end-user interaction such as a click or impression exists within a public digital Out-of-Home advertising environment, the article presents a probabilistic attribution framework for linking offline advertisement exposures to observable end-user digital behavior through defined geographical regions of exposure. Using DSR methodology, we construct and validate a spatial-temporal modeling framework that utilizes geolocation signals, sensor data harvested from devices of subjects and privacy-aware inference algorithms. Within this framework, a probabilistic viewability fence concept introduces spatial and temporal constraints on the inferred exposure while employing quality filters including dwell time, device orientation, and movement patterns. Comparative validation with attribution modeling against benchmarks set by location-based and machine learning models shows that multi-dimensional probabilistic exposure inference is applicable and effective. The framework thus turns DOOH into an accountable advertising medium, as opposed to the pure brand building medium, allowing cross-channel comparison of campaigns and data-driven actions by marketers. This article contributes to the probabilistic attribution theory for non-interactive environments and gives a practical architecture to connect sample-based behavior offline and online.

**Keywords:** *Cross-Channel Measurement, Design Science Research, Digital Out-Of-Home Advertising, Privacy-Preserving Analytics, Probabilistic Attribution, Spatial-Temporal Modeling*

## 1. Introduction

### 1.1 Background and Context

Online advertising has become more complex due to the ability of advertisers to measure click-through rates (CTR), conversion rates, and user engagement. As a result, real-time bidding for online display advertising has enabled the processing of millions of ad requests a day. The iPinYou benchmark dataset provides some insights into the sheer scale of programmatic display advertising [1] but Digital Out of Home (DOOH), such as digital billboards, digital retail store displays, digital transit displays, and displays in venues, are shown in physical environments where one-to-one granularity may not be achievable. This historical measurement

gap has kept DOOH mostly as a brand awareness advertising channel, with limited use in outcome or performance marketing.

DOOH networks are common on transport infrastructure and in public and commercial premises, and the lift and escalator market is an expanding venue for DOOH, with further research active into how passengers might interact while in transit [2]. While display infrastructure has improved, such as programmatic platforms for real-time ad serving to thousands of physical ad spaces, the challenge of establishing causal relationships between in-person ad impressions and consumer behaviors remains a methodological issue that requires advances in integrating spatial, temporal, and behavioral data across scales.

---

*Independent Researcher, USA*

## 1.2 Problem Statement and Research Gap

The main obstacle in DOOH attribution is its non-interactivity. Unlike in other channels, the ad is not directly joined to user activity. Instead, proximity to the device, the time that the viewer watched the ad in relation to its projected playback time, and behaviors indicating viewing are used to indicate engagement. Existing methods of measuring physical exposure are often based on distance methods, which are generally prone to false positives, and do not account for the multidimensionality of exposure. However, there are three major gaps that exist in the existing literature on DOOH attribution: First, systematic methods of probabilistic exposure inference, which allow for a heterogeneous mix of data sources to be integrated smoothly, are still in their infancy. Second, privacy-preserving architectures that are capable of doing attribution without disclosing who was responsible must be formalized further. Third, experiments are needed to assess the accuracy of attribution in non-interactive environments.

## 1.3 Research Objectives and Contribution

This paper thus proposes a probabilistic attribution framework for DOOH which addresses the shortcomings identified through the contributions of this section. The framework is anchored in a formal probabilistic viewability fence model which defines a multi-dimensional exposure inference criterion in terms of spatial proximity, temporal alignment and behavioral signal quality. A privacy-preserving architecture for large-scale attribution upholds analytical value while meeting data privacy requirements. The design science research approach defines design principles, artifact characteristics, and an evaluation protocol to provide guidance for constructing an attribution system. Comparison to established benchmarks in related fields shows feasibility and expected performance characteristics.

The remainder of this article is organized as follows: related work and theory are covered in

Section 2, methodology is described in Section 3, and the architecture of the framework is described in Section 4. Implementation considerations are described in Section 5, validation results and discussions are presented in Section 6, and limitations and future work are reviewed in Section 7. Section 8 concludes with implications.

## 2. Related Work and Theoretical Foundations

### 2.1 Attribution Modeling In Digital Advertising

Attribution modeling has evolved in the context of digital advertising to include multi-touch attribution methods. Dual-attention mechanism based multi-touch attribution modeling approaches are a recent advancement to capture the credit distribution over advertising touchpoints. This is achieved through usage of deep learning frameworks capturing diverse user journeys across channels and devices [3]. Probabilistic attribution approaches rely on statistical modeling to assign conversion credit in the absence of deterministic matching using machine learning algorithms to estimate exposure-to-conversion probabilities.

The deterministic-to-probabilistic transition was driven by an understanding of the consumer journey across multiple channels and devices, and the need for inference-based methods that are fundamentally strong to incomplete observability. Cross-device attribution studies have shown that probabilistic attribution can achieve strong matching accuracy for household-level attribution and set a precedent for the use of inference-based methods in a complex context.

### 2.2 Location-Based Analytics and Geospatial Inference

To improve DOOH attribution, geospatial analytics leveraging mobile device location signals is used. Trajectory data mining is the process of extracting useful patterns from

trajectory data for various applications in fields such as transportation, urban planning, and location-based services [4]. Tools have been developed in this literature to analyze raw location observations to extract behavioral patterns such as points of interest or spatiotemporal activity patterns within the data. Recent advances in LBS have continued to aim to overcome issues of accuracy, privacy, and context-based offerings indoors [8]. However, consideration for geofencing, a virtual fence that can trigger actions based on a device's location, has mainly been studied in mobile marketing. Most geofence use cases use a circular or polygonal geofence with binary in or out logic. This is limiting for DOOH attribution because this logic fails to account for the probabilistic nature of viewability of advertising, which is affected by environmental conditions, view angles, and viewer attention.

### **2.3 Privacy-Preserving Data Processing**

Privacy is a meaningful consideration for attribution systems, particularly for applications such as processing location data in public spaces. Secure aggregation protocols provide privacy-preserving machine learning, making it possible to cryptographically perform computations over globally shared datasets and outputs which do not reveal individual computation results [5]. These techniques provide mathematical guarantees against individual identification, while still providing the aggregate analysis that advertising measurement requires.

Recent audits examining existing operational personal data processing systems identified recurring compliance gaps and other architectural requirements for lawful processing under privacy regulations [14]. A set of privacy models, starting with k-anonymity, ensures that the individual records cannot be distinguished from at least k-1 other records with any combination of quasi-identifiers. When data sources are provided by third parties (such as ad networks, location data providers and measurement providers), then

a privacy-preserving architecture is required to deal with the DOOH attribution problems in practice.

## **3. Research Methodology**

### **3.1 Design Science Research Framework**

The research adopted a Design Science Research (DSR) methodology. DSR frameworks specify activities for the development of information systems artifacts [6]. DSR stage-based process models include problem identification and motivation, objectives for a solution, design and development, demonstration, evaluation and communication activities. The iterative process of DSR allows for cycles, where the results of the evaluation phase can lead to modifications of the design to improve the artifact.

The research problem of enabling measurable attribution for non-interactive DOOH advertising was identified via practitioner engagement and literature review. The solution objectives of the artifact were established and correspond to practitioner requirements and academic knowledge gaps. Specifically, they are the ability to enable probabilistic exposure inference from heterogeneous data sources, balance individual privacy with aggregate analytical utility, outperform existing proximity-based methods, and achieve scalable implementation within heterogeneous DOOH inventory.

### **3.2 Artifact Development Process**

The design process of the probabilistic attribution framework followed the principles of iterative design and consulted with experts in advertising technology, geospatial analytics, and privacy engineering. Information systems design science research stresses the need to balance utility and rigor and to solve practical problems while advancing theory [7]. Design decisions were guided mainly by established principles of probability theory, spatial statistics, and information security, with each version evaluated against a list of objectives.

The study includes a formal probabilistic viewability fence model for inferring exposure probabilities, an architectural design for privacy-preserving attribution processing, and operational physical implementation guidelines. Together, these three artifacts contribute to both the theoretical understanding of exposure inference in the context of DOOH measurement systems and provide guidance for practical real-world implementations.

### 3.3 Evaluation Methodology

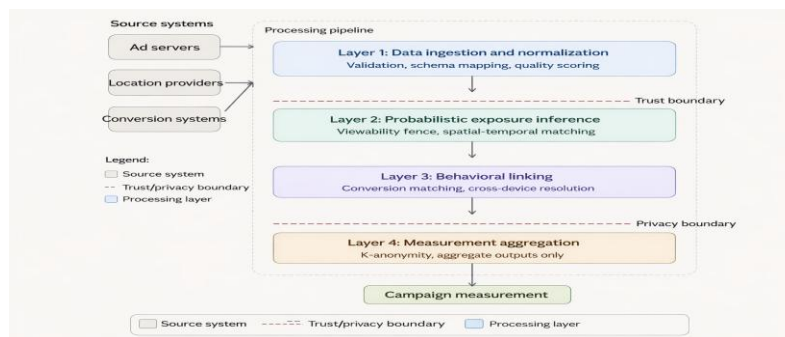
To assess the artifact, baselines were chosen from attribution models, location-based services and machine learning applications to understand feasibility and expected performance. In light of challenges with using ad effectiveness measurement in prior work, where even the most successful advertising campaigns explain only tiny fractions of variance in consumer behavior [12], the evaluation is conducted with metrics that reflect practically meaningful improvements in attribution precision and false positives.

## 4. Probabilistic Attribution Framework

### 4.1 Conceptual Architecture Overview

The proposed system treats DOOH attribution as a probabilistic inference problem, where the probability of exposure is estimated based on observable signals which are taken as proxies for the unobservable visual attention signal. Location intelligence and mobile sensing research has shown that signals derived from mobile devices characterize user behavior and social context [8]. It consists of four layers: data ingestion and normalization, probabilistic exposure inference, behavioral linking, and measurement aggregation.

Figure 1 shows the architecture of the framework, from source systems through various processing layers to the measurements. Via modularity, individual layer components can evolve without breaking interface compatibility. They are separated by trust boundaries - separated trust domains with differing levels of differential privacy - and protect per-individual data in processing enclaves while offering campaign measurement and campaign optimization capabilities.



**[Figure 1: Probabilistic Attribution Framework Architecture] The four horizontal layers each correspond to a functional area of data processing: (1) data ingestion layer, (2) exposure inference layer, (3) behavioral linking layer, (4) measurement aggregation layer. Data flows in a vertical dimension across different trust boundaries. The ingestion layer reads from source systems, including ad servers, location providers, and conversion systems. Interfaces define the connections among processing components. Privacy boundaries separate processing done at individual and aggregate levels.**

## 4.2 Probabilistic Viewability Fence Model

A probabilistic viewability fence generalizes the geofenced area by reducing multiple dimensions to a single probability of exposure. Machine learning and deep learning, effective pattern recognition and attribution techniques, are widely used in multiple domains to perform inference with limited information [9]. Formally, the exposure probability  $P(E|d,a)$  given a device observation  $d$  and an ad playback event  $a$  is equal to the product of the spatial, temporal and attentional proximity probabilities.

**Spatial Probability Component.** This probability is a function of distance from the ad placement, applying a parameterized decay function. This viewability fence defines the inner and outer radii with high, moderate, and negligible probability of being viewed. The decay parameters vary according to the dimensions of the display canvas, the viewing angle, and the presence of clutter and other visual stimuli.

**Temporal Probability Component.** Temporal probability requires an ad to be synchronous with the device's presence. In this case, temporal probability is the ratio of dwell time during which the ad was shown. Minimum dwell thresholds filter out brief observations of a device when it is likely not an exposure opportunity.

**Attention Probability Component.** The device sensors are used to sense the orientation, motion and application state. The use of mobile device sensing for environmental monitoring has been the topic of research. Accelerometer and orientation data can, for example, be used to

characterize user physical activity states [15]. For example, an user is less likely to focus visual attention on the environment when also using a mobile screen on the move.

## 4.3 Data Integration Architecture

Since the components of the data come from heterogeneous sources and different schemata, latencies, and quality levels, the data integration architecture must provide ingestion interfaces, temporal alignment and quality scoring.

**Ad Playback Data.** Playback logs from DOOH ad servers show when and how long an ad played, the creative ID, and where. Data is transferred with latencies from near-real-time to batch, which requires temporal buffering in order to be synchronized with location signals.

**Location Signal Data:** Mobile device location observations are collected from several providers through a range of means. Each observation contains an anonymized device identifier, coordinates, timestamp, and estimated accuracy, and varying quality from provider and collection context requires an assessment of data quality for any given observation.

**Conversion Data:** Downstream behavioral data from tracked events (website visits, application responses, and transaction records). Conversion data must be matched probabilistically with identifiers from the location signal, introducing uncertainty that must be accounted for in attribution estimates.

Table 1 summarizes data source characteristics and integration considerations essential for robust attribution processing.

Data Source	Typical Latency	Quality Variability	Integration Considerations
Ad playback logs	Minutes to hours	Low (structured system logs)	Temporal buffering for alignment
GPS location signals	Near real-time	Moderate (accuracy 5-20m)	Coordinate normalization, accuracy filtering
Network-based location	Near real-time	High (accuracy 50-500m)	Quality scoring, probabilistic weighting

Web conversion events	Near real-time	Low (deterministic tracking)	Identifier matching, attribution windowing
App engagement events	Near real-time	Low (SDK instrumentation)	Cross-device resolution
Transaction records	Hours to days	Moderate (POS integration variance)	Temporal alignment, partial matching

**Table 1:** Data Source Characteristics and Integration Requirements

#### 4.4 Privacy-Preserving Processing Architecture

This framework for privacy preservation depends on architectural mechanisms that limit the amount of data that an individual unit has access to, while retaining useful information at the aggregate level, and thus providing end-to-end privacy.

**Identifier Abstraction.** Irreversibly replaces device identifiers that are ingested with session-scoped tokens. Tokens are designed to evolve over time to prevent long-term tracking, but maintain session-level consistency for attribution computations.

**Processing Isolation.** Personal-level processing is done in secure enclaves; data egress is limited. Secure computation principles such as secure aggregation are used to hide individual contributions [5]. Inference computations are performed on a per-observation basis, and only the aggregate results are exposed to entities outside of the trusted boundary, not exposure-level information.

**Output Aggregation.** All outputs exposed externally are aggregates across minimum population thresholds. Aggregation provides k-anonymity: when measurement reports are released, individual users cannot be identified or characterized. User metrics have thresholds that can be set based on regulatory obligations, and the use case's sensitivity.

### 5. Implementation Considerations

#### 5.1 Scalability Architecture

During production, millions of daily location observations and thousands of ad placements

need to be processed. Our architecture consists of distributed stream processing for real-time data ingestion and batch processing for historical data analysis and model training.

**Stream Processing Layer.** Incoming data streams are validated, normalized, and quality scored by horizontally scalable processing nodes. Hierarchical spatial indexing structures like quadtrees can efficiently link geo-located observations to nearby ad locations [10]. Quadtree decomposition recursively partitions a space into quadrants that can be queried in logarithmic time, which can be used to make proximity matching operations used to infer exposures more efficient.

**Batch Processing Layer.** Attribution model calibrations, accuracy analyzes, and historical reprocessings are run through batch workflows. This partitioned storage allows for temporal and geographic parallelism. Scheduled batch jobs update the model parameters, which stream processing nodes receive to ensure similar inference behavior across nodes.

**Storage Architecture.** Required time-series optimized storage for the time-centric nature of observation data and basic range queries to manage attribution windows. Columnar formats, which allow retrieval of subsets of attributes, reduce I/O overhead for analytical workloads with a limited attribute subset.

#### 5.2 Vendor-Agnostic Integration

The architecture offers support for integrating data from diverse data sources via abstraction layers that mediate vendor-specific data formats and semantics. Data volume continues to grow and the need for cross-organizational data mining

and data integration is expected to continue to grow as well [11]. Standardized ingestion interfaces define required/optional attributes and transformation specifications for common vendor-specific ingestion formats.

Integration onboarding involves schema mapping (finding the correspondence between a vendor's attributes and the canonical schema), data quality assessment and calibration. Quality assessment checks for completeness, accuracy, and latency patterns. Calibration adjusts processing parameters based on the characteristics of vendor-specific data. Calibration improves attribution accuracy for each data source.

Table 2 presents key operational metrics and monitoring approaches derived from production deployment requirements.

Metric Category	Specific Metrics	Monitoring Approach	Alert Thresholds
Data Quality	Ingestion completeness, Schema conformance, Signal accuracy distribution	Statistical process control, Anomaly detection	>5% deviation from baseline
Processing Performance	Throughput rate, End-to-end latency, Resource utilization	Time-series monitoring, Capacity planning models	>95th percentile latency breach
Attribution Validity	Exposure rate distribution, Lift stability, Geographic consistency	Statistical validation, Cross-period comparison	Significant distribution shift
Privacy Compliance	Aggregation threshold adherence, Identifier rotation compliance	Automated audit, Configuration verification	Any threshold violation

**Table 2:** Operational Metrics and Monitoring Framework [11]

## 6. Validation and Results

### 6.1 Validation Framework And Methodology

Evaluating the effectiveness of advertising can be difficult. As with the challenge of designing validation studies in probabilistic attribution systems, research into large-scale advertising experimentation has revealed that even successful campaigns with high returns on investments only explain very small fractions of the variance in consumer behavior. Studies in the literature typically report effect sizes on the order of  $R^2 = 0.0000054$  [12]. As for many types of consumer data, the standard deviation of

### 5.3 Operational Monitoring

In production deployment, the quality of the data, performance of processing, and validity of the output is continuously monitored. Data quality metrics include completeness of ingestion, schema conformity, and anomaly detection in source feeds. Processing monitoring provides throughput, latency, and resource usage information, used for capacity planning and performance tuning.

purchases is often 10 times larger than the mean, providing large amounts of noise that make it difficult to detect treatment effects.

Experiments with fewer than 200,000 unique users may lack the power to detect meaningful effects. Power estimates for experiments with 2 million unique users are around 95% at the conventional levels of statistical importance [12]. These statistical properties justify the statistical validation of the proposed DOOH attribution framework.

In web experimentation research, when using the validation methodology, variance can be underestimated when calculating metrics that are

not independent across experimental units [13]. In a study comparing the formulaic and bootstrap estimates of variance for several of these metrics, 30% were statistically important, compared to a 5% level of importance under the null hypothesis. This suggests a systematic underestimation of variance. A bootstrap with 1000 samples can be a good variance estimator when the experimental unit is different from the randomization unit.

## 6.2 Performance Benchmarks For Attribution Models

Evaluation of attribution model architectures gives useful baselines for the performance of the system on probabilistic inference tasks. For example, dual-attention recurrent neural networks have been used to perform multi-touch conversion attribution with a large performance improvement compared to baseline methods on large-scale advertising datasets [3]. On the Criteo dataset of 700 campaigns with 16 million impressions and 45000 conversions, a classifier achieved an AUC of 0.9799 and log-loss of 0.1591. On the Miaozen dataset of 1.24 billion ad logs from 59 million users, an AUC of 0.9123 and log-loss of 0.1095 were reported.

The dual-attention attribution model outperforms state-of-the-art baselines including additive hazard models and naive probabilistic models with more than 5.5% improvement in conversion estimation performance [3]. Further experimentation on the attribution weights proves that the click-level patterns contribute far more to the final conversion attribution than the impression level patterns. The resulting distribution of weighting parameters suggests that post-click behavioral patterns are the most predictive of conversion outcomes.

These benchmarks establish new performance baselines for probabilistic attribution models trained on large-scale advertising data, showing the efficacy of learned attention as compared to hand-created credit assignment heuristics to train attribution weights via a conversion estimation objective.

## 6.3 Location Signal Processing Precision

Spatial components in the attribution framework have been validated using PDR benchmarks in location-based services literature. For example, in the domain of indoor positioning, combining WiFi fingerprinting with PDR reduced positioning error from 4.09 m to 2.32 m when employing sensor fusion via adaptive filtering of system noise [8]. These accuracy levels give perception into the expected quality of location signals for DOOH attribution applications where device location is used for exposure.

For several dimensions of signal measurements (textual data, profile, and place labeling), this inference based on incomplete signals has a coverage of 87% and an average (median) distance error of 12.2 km (4.5 km) [8]. Although these errors are too high to be directly used to infer DOOH exposure, the multi-dimensional location inference framework provides a probabilistic viewability fence where multiple dimensions of signal measurements affect the probability of being viewed.

The use of real-time Kalman filter algorithms on low-cost MEMS inertial measurement units to achieve meter-level positioning accuracy provides a baseline for the device-based exposure inference [8]. The principles of sensor fusion employed in this case can be adopted to combine the spatial proximity, temporal matching, and attention probability components of the signal to overcome the limitation of any single component.

## 6.4 Machine Learning Attribution Validation

Machine learning attribution and prediction techniques can provide further information about the validity of probabilistic systems. In the literature, oversampling techniques have been shown to considerably help model performance in prediction problems with small training datasets [9]. When applying spatial diversity oversampling, the performance of the MAE improved from 1.88 to 1.02 with the decision

tree, from 2.45 to 1.54 with MLP models, and from 1.52 to 1.36 with the SVR models.

Hybrid deep learning networks combining VAE and MLPs were found to improve upon persistence and climatology methods for forecasting tasks [9]. Thus, the model proposed here also consists of an ensemble of these networks in order to produce final predictions for the attribution task, as an ensemble of networks may produce more strong estimates than a single architecture.

These different lines of validation evidence from the attribution modeling and evaluation, the location signal processing, and the machine learning prediction support both the feasibility and the target performance of the proposed probabilistic DOOH attribution framework. While validation through a production deployment is still to be done, benchmarks enable informed expectations in terms of multi-dimensional probabilistic exposure inference performance.

Validation Area	Metric	Before	After/Value
Experiment Power	Users required for 95% power	200,000 (underpowered)	20,00,000
Variance Estimation	Metrics showing false significance	5% (expected)	30% (observed)
Attribution Model (Criteo)	AUC / Log-loss	Baseline	0.9799 / 0.1591
Attribution Model (Miaozhen)	AUC / Log-loss	Baseline	0.9123 / 0.1095
Positioning Accuracy	WiFi/PDR fusion error	4.09 m	2.32 m
ML Prediction (Decision Tree)	MAE with oversampling	1.88	1.02
ML Prediction (MLP)	MAE with oversampling	2.45	1.54
ML Prediction (SVR)	MAE with oversampling	1.52	1.36

Table 3: Validation Performance Benchmarks [3, 8, 9, 12, 13]

## 7. Limitations and Future Directions

### 7.1 Limitations

This research also has some limitations that suggest possible future avenues for research.

**Validation Scope.** Although comparisons with established benchmarks are useful, eventual deployment in real-world settings at scale introduces challenges that benchmarks cannot fully address. According to Ridder et al., there are additional deployment challenges with personal data processing systems like consent management, data minimization, and purpose limitation and in production systems, data quality variability, operational errors, and adversarial conditions may affect the attribution accuracy

beyond what can be observed in the validation set [14].

**Ground Truth Availability.** Attribution accuracy evaluation requires ground truth exposure labels, which are not directly available for non-interactive advertising. Measuring the effects of advertisements is a common challenge (and even successful campaigns explain only tiny fractions of behavioral variance [12]), but the validation of response data could additionally rely on eye-tracking or controlled exposure studies.

**Generalization Constraints.** The calibration of the framework depends on the context of deployment, such as DOOH inventory, provider characteristics, and audience behavior. Generalization to very different test conditions often requires recalibration, and may have very

different accuracy characteristics based on local environments and behavior.

Incorporation of evolving regulatory requirements and threat models. Changes in privacy regulation, the development of new or improved re-identification techniques, or shifts in social perspective regarding acceptable practices in the provision of location data may require architectural changes.

## 7.2 Future Research Directions

Future research could extend the suggested framework to overcome some of its limitations, such as utilizing computer vision signals from cameras attached to the DOOH displays themselves, which could provide direct measures of attention and, thus, improve inference. Mobile sensing literature has shown that mobile environmental sensors can capture rich contextual signals beyond location [15], offering opportunities to improve attention modeling using multi-sensor fusion [16].

Machine learning to estimate exposure probability could incorporate observed conversions to adaptively adjust models, rather than relying on fixed parameters. Cross-channel attribution models would link DOOH exposure and digital touch points for a more holistic view of the customer journey. Longitudinal studies to see how well attribution accuracy holds over time will yield more insights into how much model drift occurs and how often it needs recalibration [17,18].

## Conclusion

This paper contributes to the Digital Out-of-Home advertising attribution. The first step requires a performance metric for a passive physical medium, such as a billboard. A probabilistic viewability fence model with spatial, temporal, and attention signals is proposed as the first challenge. The second challenge is architecting a system to perform privacy-preserving analysis without sacrificing utility. The contribution also furthers Design

Science Research methodology through the development and evaluation of artifacts in attribution system research. The framework's feasibility and expected performance characteristics are shown, compared to established benchmarks of attribution models, implementations of location-based services, and machine learning algorithms. It also provides a way for DOOH to move from a brand awareness medium to a measurable input into cross-channel marketing optimization algorithms.

Furthermore, the work can be used as a groundwork for probabilistic attribution problems in non-interactive settings, and the use of Design Science Research provides a way forward for measuring systems with partially observable ground truth. Theoretically, it extends attribution modeling beyond digital venues to physical exposure inference. It provides practitioners with an architectural framework for large-scale DOOH attribution systems. The modularized structure allows to implement and to make use of certain components according to the organization's availability and the availability of data. The privacy-preserving processing architecture complies with growing regulatory requirements and can be applied to optimize advertising investment.

## References

- [1] Weinan Zhang et al., "Real-Time Bidding Benchmarking with iPinYou Dataset," arXiv, 2015. [Online]. Available: <https://arxiv.org/pdf/1407.7073>
- [2] Begüm Avdagiç et al., "The Potential of Digital Out-of-Home Advertising in the Lift Industry," 16 th Symposium on Lift & Escalator Technologies. [Online]. Available: [https://liftescalatorlibrary.org/paper\\_indexing/papers/00000574.pdf](https://liftescalatorlibrary.org/paper_indexing/papers/00000574.pdf)
- [3] Kan Ren et al., "Learning Multi-touch Conversion Attribution with Dual-attention Mechanisms for Online Advertising," ACM, 2018. [Online]. Available:

<https://dl.acm.org/doi/pdf/10.1145/3269206.3271677>

[4] YU ZHENG, "Trajectory data mining: An overview," *ACM Transactions on Intelligent Systems and Technology*, 2015. [Online]. Available:

<https://dl.acm.org/doi/pdf/10.1145/2743025>

[5] Keith Bonawitz et al., "Practical Secure Aggregation for Privacy-Preserving Machine Learning," *Proceedings of the ACM Conference on Computer and Communications Security*, 2017. [Online]. Available:

<https://dl.acm.org/doi/pdf/10.1145/3133956.3133982>

[6] Ken Peppers et al., "THE DESIGN SCIENCE RESEARCH PROCESS: A MODEL FOR PRODUCING AND PRESENTING INFORMATION SYSTEMS RESEARCH," *arXiv*. [Online]. Available:

<https://arxiv.org/pdf/2006.02763>

[7] Alan R. Hevner et al., "Design science in information systems research," *MIS Quarterly* Vol. 28 No. 1, pp. 75-105/March 2004. [Online]. Available:

[https://wise.vub.ac.be/sites/default/files/thesis\\_info/design\\_science.pdf](https://wise.vub.ac.be/sites/default/files/thesis_info/design_science.pdf)

[8] Haosheng Huang and Georg Gartner, "Current Trends and Challenges in Location-Based Services," *MDPI*, 2018. [Online]. Available: <https://www.mdpi.com/2220-9964/7/6/199>

[9] Sancho Salcedo-Sanz et al., "Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: a review," *Theoretical and Applied Climatology* (2024). [Online]. Available:

<https://link.springer.com/content/pdf/10.1007/s00704-023-04571-5.pdf>

[10] HANAN \$AMET, "The Quadtree and Related Hierarchical Data Structures," *ACM Computing Surveys (CSUR)*, Volume 16, Issue 2 (June 1984). [Online]. Available:

<https://dl.acm.org/doi/pdf/10.1145/356924.356930>

[11] David Reinse et al., "The Digitization of the World From Edge to Core," *International Data Corporation White Paper*, 2018. [Online]. Available:

<https://opis-cdn.tinkoffjournal.ru/mercury/idc-seagate-dataage-whitepaper.pdf>

[12] Randall Lewis et al., "Measuring the Effects of Advertising The Digital Frontier," *Economic Analysis of the Digital Economy*, ISBN 0-226-20684-X; 978-0-226-20684-4, 2015. [Online]. Available:

<https://www.nber.org/system/files/chapters/c12991/c12991.pdf>

[13] Thomas Crook et al., "Seven Pitfalls to Avoid when Running Controlled Experiments on the Web," *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (June 2009)* [Online]. Available:

<https://dl.acm.org/doi/pdf/10.1145/1557019.1557139>

[14] Supreeth Shastri et al., "The Seven Sins of Personal-Data Processing Systems under GDPR," *Proceedings of the USENIX Workshop on Hot Topics in Cloud Computing*, 2019. [Online]. Available:

<https://www.usenix.org/system/files/hotcloud19-paper-shastri.pdf>

[15] Prashanth Mohan et al., "Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones," *Proceedings of the 6th ACM conference on Embedded network sensor systems (November 2008)* [Online]. Available: <https://dl.acm.org/doi/pdf/10.1145/1460412.1460444>

[16] J. Boadi-Mensah, "Waste management in the 21st century: Challenges, opportunities, and sustainable solutions," *Journal of Sustainable Agriculture and Environmental Innovations*, vol. 1, no. 3, pp. 1–11, 2025.

[17] D. Joshi, "From quality to value: How robust data governance drives analytics-led business

performance,” *Journal of International Crisis and Risk Communication Research*, pp. 403–411, 2024. [Online]. Available:

<https://doi.org/10.63278/jicrcr.vi.3528>

[18] N. Fernandes, “Psychoeducational group facilitation as a tool for mental health equity in diverse populations,” *Review of Contemporary Philosophy*, vol. 21, no. 1, pp. 16–24, 2022.