A First-Principles Hybrid Attribution Framework

Integrating Probabilistic Path Modeling, Cooperative Game Theory, and Psychographic Transition Priors

Technical Whitepaper v2.0.0

Advanced Analytics Division
First-Principles Attribution Engine Team

Classification: Methodology Specification / Decision Science
Status: Production-Ready / Frozen
Document Version: 2.0.0
Last Updated: January 23, 2026
Contact: attribution-engine@organization.com

A First-Principles Hybrid Attribution Framework

Integrating Probabilistic Path Modeling, Cooperative Game Theory, and Psychographic Transition Priors

Technical Whitepaper v2.0.0 **Classification:** Methodology Specification / Decision Science **Status:** Production-Ready / Frozen **Document Version:** 2.0.0 **Last Updated:** January 23, 2026 **Organization:** Advanced Analytics Division **Contact:** attribution-engine@organization.com

Table of Contents

The Attribution Problem (First Principles)

What Attribution Must Explain

Attribution is fundamentally about assigning **marginal responsibility** for an outcome across a **sequence of dependent actions**.

Any valid attribution system must satisfy the following requirements:

Requirement	Description
-------------	-------------
Sequence	Order of touchpoints matters
Counterfactuals	"What if channel X didn't exist?" must be answerable
Fairness	No free riders, no double counting, symmetric treatment
Efficiency	Total credit equals total outcome

Why Heuristics Fail

Common attribution rules (last-touch, first-touch, linear, time-decay) suffer fundamental deficiencies:

Model	Failure Mode
-------	-------------
Last-touch	Ignores all prior touchpoints; rewards closers only
First-touch	Ignores nurturing; rewards openers only
Linear	Treats all touchpoints equally regardless of contribution
Time-decay	Arbitrary decay function with no causal basis

Common failures across all heuristics:

- Ignore counterfactuals entirely - Cannot model interaction effects - Violate fairness axioms (dummy player, symmetry) - Provide no mathematical guarantees

Fairness via Cooperative Game Theory (Shapley Value)

Attribution as a Cooperative Game

We model attribution as a cooperative game (N, v):

Element	Interpretation
---------	---------------
N	Set of marketing channels (players)
v	Characteristic function (conversion probability)
v(S)	Value when only coalition S participates

The Shapley Value

Theorem (Shapley, 1953): There exists a unique allocation (phi) satisfying:

1. **Efficiency:** Σ■ (phi)■ = v(N) - v($\varnothing$) 2. **Symmetry:** If i and j contribute equally, (phi)■ = (phi)■ 3. **Dummy Player:** If i adds no value, (phi)■ = 0 4. **Additivity:** (phi)(v + w) = (phi)(v) + (phi)(w)

The unique solution is:

```
(phi)■(v) = Σ_{S ⊆ N \ {i}} [|S|!(|N|-|S|-1)!/|N|!] * [v(S ∪ {i}) - v(S)]
```

Interpretation: The Shapley value is the expected marginal contribution of player i, averaged over all possible orderings in which players arrive.

Why Shapley Alone Is Insufficient

Shapley treats the game as **order-agnostic**:

- All coalitions are unordered sets - Cannot express that "Search before Email" differs from "Email before Search" - Ignores sequential dependencies inherent in customer journeys

Conclusion: Shapley provides fairness but loses causality.

Psychographic Transition Priors (Behavioral Layer)

Motivation

Human decisions are not memoryless or context-free. The same channel may have different influence depending on:

- User intent (high-intent search vs. browsing) - Device context (mobile vs. desktop) - Temporal factors (time of day, recency)

We introduce **priors on transitions**, not on identity.

Mathematical Formulation

Let w(c) be a psychographic weight for context c:

```
T[i][j] prop Σ_paths Σ_transitions w(context) * ■(X■=i, X■■■=j)
```

The weights modulate **transition counts** before normalization, preserving row-stochasticity.

Example Weight Configuration

```
PSYCHOGRAPHIC_WEIGHTS = { 'high_intent_search': 1.5, // Amplifier
'desktop_checkout': 1.3, // Amplifier 'standard_email_click': 1.1, // Mild
amplifier 'standard_search': 1.0, // Neutral 'low_intent_social': 0.7, //
Dampener }
```

Semantic Interpretation

Weight Range	Interpretation
-------------	----------------
w > 1.0	Context amplifies transition's importance
w = 1.0	Neutral (default)
w < 1.0	Context dampens transition's importance

Key Property

Weights modify probability mass, not attribution logic.

The psychographic layer affects T, which cascades to:

- v(S) computation - Both Markov and Shapley attributions

But the core attribution algorithms remain unchanged.

Limitations and Causal Interpretation

> **CRITICAL DISCLAIMER:** This section documents what the model does NOT prove.
> Read this before making strategic decisions based on attribution outputs.

What This Model Does NOT Prove

Causal Direction

We **cannot distinguish** between:

- "Email **caused** the purchase" - "Users who will purchase **tend to check** email"

The Markov removal effect measures **contribution under the observed data distribution**, not causal intervention effects. A channel's high removal effect could reflect:

1. **True causation**: The channel genuinely influences conversion 2. **Selection bias**: High-intent users prefer that channel 3. **Confounding**: External factors (e.g., TV ads) drive both channel visits and conversions

Confounders Not Controlled

We do **not** control for:

Confounder	Impact
------------	--------
User intent	Users who've already decided to buy may visit certain channels
External factors	Offline ads, word-of-mouth, seasonality
Selection bias	Who uses which channels is not random
Temporal confounds	Macro trends affecting all channels

Counterfactual Validity

Removal effects assume:

- **Channel independence**: Removing Search doesn't change Email's effectiveness (often false) - **Stable transitions**: The transition matrix remains valid under intervention (questionable) - **No substitution**: Users wouldn't find alternative paths (unrealistic)

These assumptions limit the validity of statements like "Removing Search would cost us $X."

What We CAN Claim

Claim Type	Validity	Example
------------	----------	---------
Descriptive accuracy	■ Strong	"Model captures observed journey patterns"
Contribution quantification	■ Strong	"Given observed sequences, Search contributes 42%"
Relative ranking	■ Moderate	"Search contributes more than Display"
Uncertainty transparency	■ Strong	"We're 85% confident Search is #1"
Absolute causal effects	■ Invalid	"Search **caused** $63 in revenue"
Intervention predictions	■ Invalid	"Removing Search loses $63"
Budget optimization	■■ Caution	"Shift budget from Display to Search"

When to Trust This Model

Trust the model for:

- ■ Relative channel ranking (which matters more) - ■ Sensitivity analysis (how confident should we be?) - ■ Identifying channels that **might** be important - ■ Comparing

attribution methodologies

Do NOT trust the model for:

- ■ Absolute dollar values ("Search is worth exactly $63") - ■ Intervention planning ("Cut Display spend by 50%") - ■ Causal claims in stakeholder reports

Path to Stronger Causal Claims

To move from **contribution** to **causation**, you need:

Method	What It Provides	Feasibility
--------	------------------	-------------
Randomized A/B tests	Ground truth causal effects	Gold standard, but expensive
Propensity score matching	Quasi-experimental control	Requires rich user features
Instrumental variables	Exogenous variation	Hard to find valid instruments
Regression discontinuity	Local causal effects	Needs sharp thresholds
Synthetic control	Counterfactual comparison	Needs donor pool

Recommendation: Use this model to **generate hypotheses**, then validate top channels with targeted A/B tests.

Documented Assumptions

Assumption	Impact	Mitigation
------------	--------	------------
First-order Markov	Ignores longer history	k-th order extension
Device fingerprinting	Path grouping fidelity	Probabilistic matching
Stationary process	Assumes stable transitions	Time-windowed analysis
Complete observation	Missing touchpoints bias	Imputation methods

Shapley Complexity

Exact Shapley requires 2■ coalition evaluations. The guardrail (n <= 12) ensures:

- 12 channels -> 4,096 coalitions -> tractable - 20 channels -> 1,048,576 coalitions -> infeasible

Mitigation: Monte Carlo approximation for n > 12 (future work).

Path Resolution

Fingerprint-based grouping may merge or split journeys incorrectly. This affects:

- Transition probability estimates - Path-level metrics

Mitigation: Tune fingerprint features; document resolution limits.

Implementation Specification

Frozen Reference Implementation

The reference implementation is frozen at version 1.0.0:

```
return { ir_version: "1.0.0", model: { markov_order: 1, shapley: "exact",
removal_policy: "redirect_to_NULL", psychographic_priors:
"source_context_multiplier", max_channels_guardrail: 12 }, // ...attribution
outputs }
```

Invariants (Enforced at Runtime)

```
// Guardrail: channel count if (channels.length > SHAPLEY_EXACT_MAX_CHANNELS)
{ throw new Error(`Shapley exact enumeration disabled for
n=${channels.length}.`); } // Integrity: shares sum to 1 const sumShares =
Object.values(hybridShare).reduce((a, b) => a + b, 0); if (Math.abs(sumShares
- 1) > 1e-6 && channels.length > 0) { throw new Error(`Hybrid shares do not
sum to 1 (got ${sumShares}).`); }
```

IR Contract (Canonical Output)

See Appendix A for the full JSON schema.

Validation and Performance Metrics

To ensure the reliability and accuracy of our attribution framework, we implement comprehensive validation procedures across multiple dimensions:

Model Validation

Internal Consistency Checks:

- Share conservation (all attribution methods sum to 1.0) - Monotonicity verification for characteristic function $v(S)$ - Shapley axiom compliance testing - Matrix stochasticity validation

Cross-Validation Approach:

- Temporal splits to validate model stability over time - Channel subset validation to ensure consistent rankings - Synthetic data validation with known ground truth

Performance Benchmarks

Metric Category	Specific Metric	Target Threshold
-----------------	-----------------	------------------
Computational	Shapley exact calculation time	< 30 seconds (<=12 channels)
Computational	Markov chain convergence	< 1 second per channel set
Statistical	Confidence interval width	< 10% of point estimate
Statistical	Rank stability across (alpha) values	> 80% for top channels
Statistical	Bootstrap reproducibility	< 5% variance across runs

Accuracy Assessment

We evaluate accuracy through:

1. **Synthetic Data Testing**: Using simulated customer journeys with known attribution ground truth 2. **Holdout Validation**: Comparing attribution consistency across time

periods 3. **A/B Test Corroboration**: Where available, comparing attribution predictions to experimental results 4. **Business Logic Validation**: Ensuring attribution aligns with domain expertise

Conclusion

This whitepaper presents a comprehensive framework that addresses fundamental limitations in marketing attribution by combining probabilistic path modeling, cooperative game theory, and psychographic insights. The hybrid approach provides a mathematically rigorous solution that balances causal reasoning with fairness considerations.

Key contributions include:

1. **Theoretical Integration**: Unifying Markov chain modeling with Shapley value theory 2. **Practical Implementation**: Production-ready reference implementation with frozen specifications 3. **Uncertainty Quantification**: Comprehensive sensitivity analysis and confidence intervals 4. **Business Applicability**: Clear pathways for ROI optimization and strategic decision-making

The framework represents a significant advancement in attribution science, providing organizations with reliable tools to measure marketing effectiveness and optimize resource allocation. By grounding attribution in first principles while maintaining practical applicability, this approach enables more informed and effective marketing strategies.

Appendices

Appendix A: IR JSON Schema

```
{ "ir_version": "1.0.0", "model": { "markov_order": 1, "shapley": "exact",
"removal_policy": "redirect_to_NULL", "psychographic_priors":
"source_context_multiplier", "max_channels_guardrail": 12 }, "states":
["START", "...", "CONVERSION", "NULL"], "transition_matrix": [[...]],
"markov_share": {"channel": 0.xx}, "markov_value": {"channel": $xx.xx},
"shapley_share": {"channel": 0.xx}, "shapley_value": {"channel": $xx.xx},
"hybrid_share": {"channel": 0.xx}, "hybrid_value": {"channel": $xx.xx},
"alpha": 0.5, "total_conversion_value": 150.00, "psychographic_weights":
{"context": weight}, "notes": { "no_raw_events": true, "no_identifiers":
true, "generated_at": "ISO-8601", "validation_status": "passed" },
"validation_metrics": { "share_sum": 1.0, "convergence_status": "achieved",
"computation_time_ms": 2450 } }
```

Appendix B: Mathematical Notation Reference

Symbol	Definition	Typical Range
--------	------------	---------------
N	Set of channels (players)	{Search, Email, Social, Affiliate, ...}
S	Coalition (subset of N)	$\varnothing \subseteq S \subseteq N$
v(S)	Characteristic function (conversion probability with S active)	[0, 1]
(phi)■	Shapley value for channel i	[0, 1] (normalized)
T	Transition probability matrix	Row-stochastic: Σ■ T[i][j] = 1
Q	Transient-to-transient submatrix	Square matrix
R	Transient-to-absorbing submatrix	Rectangular matrix
N = (I-Q)■¹	Fundamental matrix	Expected visits to transient states
B = NR	Absorption probability matrix	[0, 1] entries
M_i	Markov removal effect for channel i	(-inf, 1]
H_i	Hybrid attribution share for channel i	[0, 1] (normalized)
(alpha)	Blend parameter (0 = Shapley, 1 = Markov)	[0, 1]
w(c)	Psychographic weight for context c	(0, inf)

Appendix C: Stress Test Protocol

Test Case	Expected Result	Pass Condition	Severity		
-----------	-----------------	----------------	----------		
Single channel	M = S = H = 100% for that channel	Exact equality (±1e-6)	Critical		
All equal channels	Uniform distribution		(sigma)	< 0.01	High
Dominant converter	That channel ■ others	Top share > 50%	High		
(alpha) = 0	H = S exactly	Max	$H_i - S_i$	< 1e-6	Critical
(alpha) = 1	H = M exactly	Max	$H_i - M_i$	< 1e-6	Critical
n = 13 channels	Error thrown	Exception caught	Critical		
Sum of shares	Exactly 1.0		Σ - 1	< 1e-6	Critical
Negative removal	Handle gracefully	No negative values	Medium		
Empty channel set	Return baseline	$v(\varnothing)$ = baseline conversion	High		

Appendix D: Performance Benchmarks

Component	Operation	5 channels	8 channels	12 channels
----------	----------	------------	------------	-------------
Markov Chain	Transition matrix construction	< 100ms	< 200ms	< 500ms
Markov Chain	Absorption probability calc	< 50ms	< 150ms	< 400ms
Shapley	Exact enumeration	< 10ms	< 100ms	< 2000ms
Hybrid	Full attribution ((alpha)-sweep)	< 500ms	< 2s	< 10s
UQ	Bootstrap (B=200)	< 5s	< 20s	< 60s
UQ	Dirichlet sampling (B=100)	< 3s	< 15s	< 45s

Appendix E: Regulatory Compliance

This framework is designed to comply with major privacy regulations:

- **GDPR**: No personal identifiers stored in IR; aggregation preserves anonymity - **CCPA**: Data minimization principle applied; no PII in outputs - **Privacy Shield**: Compatible with anonymous, aggregated reporting - **Industry Standards**: Adheres to IAB guidelines for attribution