

# A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

Sudipkumar Ghanvat<sup>1</sup>, Shreya Joshi<sup>2</sup>, Rohan Jamdagni<sup>3</sup>, Aditi Shintre<sup>4\*</sup>

<sup>1</sup>Sr. Director & Head - Data & AI, VRIO Digital Dallas, USA

<sup>2</sup>Sr. Data Analyst, American Airlines, Texas, USA

<sup>3</sup>Principal Solution Architect, Zeptta, Texas, USA

<sup>4</sup>Research Engineer, Neowesolutize Technology Pvt Ltd Pune, Maharashtra, India.

\*Corresponding Author.

## OPEN ACCESS

### Article Citation:

Sudipkumar Ghanvat<sup>1</sup>, Shreya Joshi<sup>2</sup>, Rohan Jamdagni<sup>3</sup>, Aditi Shintre<sup>4\*</sup>, "A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys", International Journal of Recent Trends in Multidisciplinary Research, September-October 2026, Vol 6(05), 127-136.



©2026 The Author(s). This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Published by 5<sup>th</sup> Dimension Research Publication

**Abstract:** Digital marketing campaigns generate increasingly complex, multi-channel customer journeys, yet conventional multi-touch attribution (MTA) approaches, including heuristic rules, Markov chain models, game-theoretic methods, and deep-learning approaches—struggle to represent the cumulative, delayed, and state-dependent effects that characterize customer behavior. This article proposes a non-Markovian Dynamic Bayesian Network (DBN) framework for MTA that represents customer journeys as sequences of observed touchpoints generated by latent behavioral (engagement/intent) states and extends the conventional first-order DBN by allowing the latent state at time  $t$  to depend on multiple preceding latent states. Channel-level attribution is computed through counterfactual removal-effect analysis on the resulting latent-state dynamics, enabling delayed and cumulative channel effects to be incorporated into attribution. The framework is positioned relative to heuristic, Markov-chain, Shapley-value, deep-learning, and graphical point-process approaches, and an experimental design combining synthetic non-Markovian simulations with real customer-journey data is proposed for evaluating attribution accuracy, predictive performance, and computational cost. An illustrative Monte Carlo simulation demonstrates the mechanism: the Email removal effect increases from 0.04207 under a first-order specification to 0.04863 under a second-order specification, representing an approximately 15.6% increase when cumulative repeated-exposure effects are introduced. The proposed framework provides an interpretable probabilistic alternative to black-box sequence models while extending the temporal depth of latent-state attribution beyond conventional first-order DBNs. Future empirical evaluation is required to establish its effectiveness, scalability, and generalizability across real-world customer journeys.

**Keywords:** multi-touch attribution; dynamic Bayesian networks; non-Markovian models; customer journey analytics; explainable AI; marketing analytics

## 1. Introduction

Contemporary digital marketing campaigns unfold across an increasingly dense set of channels—social media, search engines, email marketing, display advertising, and mobile applications—through which customers encounter a firm repeatedly before completing a purchase or other conversion event. Understanding how much credit each of these touchpoints deserves for an eventual conversion is a foundational problem in marketing analytics, because attribution outcomes directly shape how marketing budgets are allocated across channels<sup>1</sup>.

Early approaches to this problem relied on simple heuristics—first-touch, last-touch, and linear attribution—that assign

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

credit according to a fixed rule rather than on the basis of observed data. These rules are easy to compute and explain, but they implicitly assume that a single touchpoint, or an equal share across all touchpoints, fully determines conversion, an assumption long recognized as biased and oversimplified<sup>2</sup>. Data-driven alternatives emerged to address this limitation. Markov chain attribution models represent the customer journey as a sequence of channel “states” and estimate each channel's contribution through a removal-effect calculation—the predicted drop in conversion probability when a channel is removed from the transition graph<sup>2</sup>. Building on cooperative game theory, Shapley-value and related bankruptcy-based attribution rules instead treat marketing channels as players in a cooperative game and allocate credit according to axiomatic fairness properties such as efficiency, symmetry, and the null-player property<sup>3</sup>. More recently, deep-learning architectures—including temporal convolutional networks combined with layer-wise relevance propagation—have been proposed to capture long and complex B2B customer journeys with many stage transitions<sup>4</sup>, and graphical point-process frameworks have been developed to model the full relational structure among touchpoint types and to compute path-level removal effects<sup>5</sup>.

Despite this progression, a central limitation persists across nearly all of these approaches: the Markov property. Markov chain attribution models, by construction, assume that the probability of the next touchpoint—or of conversion—depends only on the current state and not on the sequence of states that preceded it. This memoryless assumption is computationally convenient but is at odds with substantial evidence that customer behavior exhibits cumulative, delayed, and state-dependent effects: prior exposures shape a customer's underlying intent or engagement level in ways that persist and compound over multiple subsequent interactions<sup>6,7</sup>. Deep-learning sequence models can, in principle, learn such long-range dependencies, but they do so as opaque, high-dimensional functions whose internal representations are difficult to translate into the kind of explanation a marketing manager can act upon<sup>8,9</sup>.

This creates a tension that the present article seeks to resolve existing attribution models typically trade off temporal richness against interpretability. Markov chain and Shapley-value models are interpretable but largely memoryless; deep-learning and graphical point-process models can represent richer temporal and relational structure but are comparatively opaque or computationally demanding<sup>5</sup>. Dynamic Bayesian Networks (DBNs)—probabilistic graphical models that explicitly represent how observed and latent variables evolve over time—offer a potential middle ground. DBNs have already been used to model latent customer states underlying sequential behavior in marketing and health contexts, including multistage shopping journeys<sup>6,10</sup>, clickstream-based shopping-phase detection and longitudinal engagement dynamics<sup>11</sup>. However, standard DBNs retain a first-order Markov assumption on latent-state transitions, which limits their ability to represent the long-memory dependencies documented in customer journey research.

This article proposes a Dynamic Bayesian Network framework for multi-touch attribution that relaxes the first-order Markov assumption on latent states. In the proposed model, each observed touchpoint is generated conditionally on an underlying latent behavioral state, but—unlike a standard DBN—the latent state at time  $t$  is permitted to depend on multiple preceding latent states rather than only on the immediately preceding one. Channel-level attribution is then computed through counterfactual removal-effect analysis performed on the fitted non-Markovian latent-state dynamics, in the spirit of removal-effect attribution but operating on inferred behavioral states rather than directly on observed touchpoint sequences<sup>2,5</sup>.

The proposed framework makes three contributions to the multi-touch attribution literature. First, it introduces a Dynamic Bayesian Network architecture for multi-touch attribution that links observed touchpoints to latent customer engagement/intent states. Second, it extends this architecture with a non-Markovian, long-memory dependency structure among latent states, enabling the model to represent cumulative and delayed marketing effects that first-order Markov and hidden Markov model (HMM)-based approaches cannot capture. Third, it offers an interpretable probabilistic alternative to black-box deep-learning attribution models, retaining the transparency associated with explicit transition and emission probabilities while extending temporal depth beyond what conventional DBNs allow<sup>8</sup>.

The remainder of the article is organized as follows. Section II reviews the literature on multi-touch attribution, Dynamic Bayesian Networks and latent-state behavioral models, non-Markovian sequence modeling, and interpretability in marketing analytics, research gap and positions the proposed framework relative to existing approaches. Section III presents the formal specification of the non-Markovian DBN, its inference and learning procedures, and the counterfactual attribution mechanism. Section IV describes experimental design, including data sources, baseline models, and evaluation metrics. Section V presents a small illustrative simulation. Section VI discusses expected results and positioning relative to baselines, Section VII discusses managerial implications, Section VIII discusses limitations and future research directions, and Section IX presents the conclusion of the article.

## 2. Literature Review

### From Heuristics to Algorithmic Multi-Touch Attribution

Heuristic attribution rules—first-touch, last-touch, linear, time-decay, and position-based models—assign conversion credit according to a predetermined rule applied to a touchpoint's position within the journey, independent of the actual statistical relationship between that touchpoint and conversion<sup>2</sup>. Their main advantages are simplicity and transparency; their main weakness is that the assigned credit need not correspond to the touchpoint's true contribution.

Markov chain attribution models replace these fixed rules with a transition-probability model estimated from observed journey data, in which each unique touchpoint—together with start, conversion, and non-conversion states—is treated as a node in a directed graph. Channel-level contribution is computed via the removal effect: a channel's transition edges are removed from the graph, the conversion probability is recomputed, and the resulting drop in conversion probability is attributed to that channel<sup>2</sup>. This approach captures the sequential, path-dependent nature of customer journeys more faithfully than heuristic rules, but it inherits the Markov property—the assumption that the next state depends only on the current state.

A parallel line of work has approached attribution through cooperative game theory. Molina et al.<sup>3</sup> formalize the multi-

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

touch attribution problem as a transferable-utility game in which marketing channels are treated as players who jointly produce a conversion “payoff,” and propose the Shapley value—together with bankruptcy-problem-based rules such as the constrained equal-losses rule—as attribution mechanisms that satisfy axiomatic fairness properties. These models can incorporate the order and repetition of channels within a path, but the characteristic function defining each channel's marginal contribution is typically derived from an underlying probabilistic, often Markovian, model of the journey, so the memoryless limitation can persist even within a game-theoretic attribution layer. Gaure et al.<sup>2</sup> provides a complementary comparison of heuristic and ensemble-based attribution approaches, underscoring that the choice of underlying model substantially affects which channels are favored once attribution moves beyond simple rules.

Machine-learning approaches have sought to relax these structural assumptions by learning the mapping from touchpoint sequences to conversion directly from data. Agrawal et al.<sup>4</sup> propose a deep-learning framework for B2B customer journeys in which each stage transition in the customer funnel is modeled by a Temporal Convolutional Network (TCN), with a global Stage-TCN conversion model combining these stage-level models, and channel- and touchpoint-level attribution computed via layer-wise relevance propagation. Such models can represent long-range dependencies and complex non-linear interactions among touchpoints, but the resulting attribution values are derived from relevance scores propagated through a deep network rather than from an explicit, human-interpretable probabilistic structure<sup>8</sup>.

Most recently, Tao et al.<sup>5</sup> proposed a graphical point-process framework that represents customer-level path-to-purchase data as a set of asynchronous event streams and learns the full relational structure among touchpoint types using a temporal point-process model. Building on this graphical structure, they define graphical attribution methods that compute path-level conversion credit using a probabilistic generalization of the removal effect. Conceptually, this framework is the closest recent precedent to the present proposal in that it explicitly moves beyond pairwise, adjacent-state transitions toward a representation of how the entire prior history of touchpoints jointly influences conversion. However, it represents this history through the relational structure of observed touchpoint types and their mutually exciting effects, rather than through an explicit layer of latent behavioral or intent states—the mechanism through which the present article's non-Markovian DBN represents long-memory effects.

### Dynamic Bayesian Networks and Latent-State Models of Customer Behavior

A Dynamic Bayesian Network is a probabilistic graphical model that extends a Bayesian network across discrete time steps, typically represented as a “two-time-slice” structure in which variables at time  $t$  are conditioned on variables at times  $t$  and  $t - 1$ <sup>12</sup>. DBNs and their close relative, the hidden Markov model (HMM), have a long history of use in marketing for representing unobservable customer states—such as relationship stage, engagement level, or shopping intent—that evolve over time and probabilistically generate observable behaviors such as purchases, visits, or clicks.

Recent work has extended this tradition to multistage digital shopping journeys. Wei et al.<sup>10</sup> propose a multistage Dynamic Bayesian Network (MS-DBN) for product recommendation that explicitly adopts a marketing-funnel perspective: the model incorporates two latent layers to capture variability in consumers' interest shifts across funnel stages and includes identification strategies that dynamically detect otherwise-invisible stage transitions and interest shifts, demonstrating superior performance over baseline recommendation approaches on large-scale real-world data. This work establishes that DBN architectures with multiple latent layers can recover psychologically meaningful, stage-structured representations of consumer journeys—precisely the kind of representation a DBN-based attribution model needs in order to support counterfactual reasoning about channel contributions at different funnel stages.

In e-commerce clickstream analytics, Ozyurt et al.<sup>6</sup> develop ClickstreamDMM, an attentive deep Markov model that combines an attention mechanism for capturing long-term dependencies in clickstream sequences with a latent-variable model representing different shopping phases (goal-directed search versus browsing), in order to predict the risk of session exit without purchase. Related work by Hatt and Feuerriegel<sup>7</sup> develops a duration-dependent hidden Markov model for detecting user exits, explicitly addressing a limitation of standard HMMs: that the expected duration of a latent state does not depend on how long the user has already spent in that state. Both lines of work demonstrate that incorporating richer temporal structure into latent-state models meaningfully improves predictive performance over standard first-order HMMs in marketing-relevant tasks, providing empirical motivation for relaxing the Markov assumption in the attribution setting as well.

Outside the marketing domain, De La Torre et al.<sup>11</sup> apply a DBN to model the dynamic relationship between user engagement and physical-activity behavior over a year-long mobile-health intervention, illustrating that DBNs remain a viable and interpretable tool for modeling longitudinal behavior-change dynamics at scale, including settings with intermittent and irregularly timed interactions—conditions that closely resemble the irregular, asynchronous touchpoint sequences observed in digital marketing journeys.

From a methodological standpoint, Shiguihara et al.<sup>12</sup> provide a systematic survey of DBN modeling, learning, and inference techniques, summarizing structure-learning, parameter-learning, and inference algorithms—including exact inference, sampling-based methods, and approximate or variational approaches—across application domains drawn predominantly from high-quality journals. Drury and Smith<sup>13</sup> further discuss DBN elicitation and structure specification in settings where direct estimation from data must be complemented by expert input—a consideration likely to be relevant when specifying the number and interpretation of latent engagement states in a marketing-attribution DBN, where ground-truth latent states are, by definition, unobserved.

### Beyond First-Order Markov: Non-Markovian and Long-Memory Sequence Representations

The defining limitation motivating this article is the first-order Markov assumption embedded in conventional Markov chain attribution models, standard HMMs, and conventional DBNs alike: the next state—whether an observed touchpoint or a

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

latent behavioral state—is assumed to depend only on the current state, not on the broader history of prior states<sup>7</sup>. Several strands of research have sought to relax this assumption from different directions. Duration-dependent and semi-Markov extensions relax the assumption that state durations are memoryless, allowing latent states to become “sticky” and thereby implicitly encoding a limited form of history-dependence<sup>6,7</sup>. Higher-order Markov chains extend the dependency window from one to  $k$  previous states, at the cost of an exponentially growing state space.

The graphical point-process framework of Tao et al.<sup>5</sup> represents a further departure from first-order dependence: rather than modeling transitions between adjacent states, it models the full relational structure—a directed graph of excitation effects—among all touchpoint types observed across a path, so that a touchpoint occurring early in a journey can be modeled as directly influencing the conversion intensity much later, independent of intervening touchpoints. This is, in effect, a non-Markovian representation, but it operates at the level of observed event types and their mutual excitation rather than via an explicit latent-state layer.

The framework proposed in this article pursues long-memory representation through a third route: rather than extending the dependency window of observed transitions (higher-order Markov chains) or modeling relational excitation among observed event types (graphical point processes), it extends the dependency structure of the latent behavioral-state sequence itself, allowing the latent state at time  $t$  to depend on the latent states at times  $t - 1$  through  $t - p$ . This keeps the observation model—touchpoint generation conditional on latent state—unchanged and computationally tractable, while concentrating the long-memory, cumulative-effects structure in the latent layer, where, as Wei et al.<sup>10</sup> and Ozyurt et al.<sup>6</sup> demonstrate, it can be given a substantively meaningful interpretation in terms of evolving customer engagement or intent.

### Interpretability and Explainable AI in Marketing Analytics

As deep-learning and other black-box methods have entered marketing analytics, a substantial literature has developed around explainable AI (XAI) and interpretable machine learning. Marcinkevics and Vogt<sup>8</sup> provide a methods-centric overview of interpretable and explainable machine learning, distinguishing inherently interpretable model classes from post-hoc explanation methods applied to opaque models, and arguing that the inductive biases built into a model's structure substantially shape what kinds of explanations are even possible. Ali et al.<sup>9</sup> similarly synthesize the state of XAI research and identify a persistent gap between technically sophisticated explanation methods and the practical requirement for trustworthy, decision-relevant explanations in applied domains.

For multi-touch attribution specifically, this distinction matters because the “explanation” a marketing manager needs—how much a channel's budget should change, and why—is most naturally expressed in terms of a model's own internal probabilistic structure (transition probabilities, latent-state occupancy, emission probabilities) rather than via post-hoc approximations of a deep network's decision boundary. DBNs belong to the class of inherently interpretable models in the taxonomy of Marcinkevics and Vogt<sup>8</sup>: their parameters have direct probabilistic meaning, and counterfactual quantities such as removal effects can be computed by directly manipulating the model's conditional probability tables rather than by approximating gradients or surrogate models. This is the principal interpretability advantage motivating the proposed non-Markovian DBN as an alternative to deep-learning attribution architectures such as that of Agrawal et al.<sup>4</sup>.

### Customer Journey Complexity and the Omnichannel Context

The practical urgency of richer attribution models is reinforced by research on the increasing complexity of customer journeys themselves. Gao et al.<sup>14</sup> show that the way firms integrate channels shapes the customer experience across the journey, with implications for how and how cumulatively exposures across channels affect downstream outcomes. Vollrath and Villegas<sup>1</sup> argue that digital marketing analytics should be organized around the customer decision journey as a strategic framework rather than around isolated channel-level metrics, cautioning against an “analytics myopia” that optimizes individual touchpoints without regard to their role in a longer journey. Mahajan et al.<sup>15</sup>, in a systematic review of marketing analytics and consumer behavior research, similarly identify the modeling of complex, multi-stage consumer journeys as an area where existing analytical techniques remain underdeveloped relative to the complexity of the phenomena being studied.

Taken together, this body of work indicates that (a) customer journeys are becoming longer, more cross-channel, and more cumulative in their effects; (b) existing attribution methods either assume away this cumulative structure (heuristics, Markov chains, Shapley-value models), capture it at the cost of interpretability (deep learning), or represent it at the observed-event rather than latent-behavioral level (graphical point processes); and (c) DBN-based latent-state models have already proven effective at representing psychologically meaningful customer states in adjacent marketing applications, but typically retain a first-order Markov structure on those latent states. Table 1 summarizes this positioning.

| Approach                                    | Representative References | Interpretability | Temporal Depth / Memory | Captures Latent Intent States | Relative Computational Cost |
|---------------------------------------------|---------------------------|------------------|-------------------------|-------------------------------|-----------------------------|
| Heuristic (first-touch, last-touch, linear) | [2]                       | High             | None                    | No                            | Very low                    |
| Markov chain (removal effect)               | [2]                       | High             | First-order             | No                            | Low                         |
| Shapley value / game-theoretic              | [3]                       | High (axiomatic) | Depends on underlying   | No                            | Moderate                    |

|                                                   |              |                    | journey model<br>(often first-order)       |                            |               |
|---------------------------------------------------|--------------|--------------------|--------------------------------------------|----------------------------|---------------|
| Deep learning<br>(TCN + relevance<br>propagation) | [4]          | Low (post-<br>hoc) | Long-range<br>(learned)                    | Implicit /<br>unidentified | High          |
| Graphical point<br>process (removal<br>effects)   | [5]          | Moderate           | Long-range<br>(relational /<br>excitation) | No                         | High          |
| DBN / multistage<br>latent-state models           | [6,10]       | High               | First-order on<br>latent state             | Yes                        | Moderate      |
| Proposed non-<br>Markovian DBN                    | This article | High               | p-th order on<br>latent state              | Yes                        | Moderate-high |

Table no 1: Comparative Positioning of Multi-Touch Attribution Modeling Approaches

The literature reviewed above suggests that attribution modeling approaches occupy different points along two dimensions that matter for managerial adoption: interpretability and temporal/relational depth. Heuristic and Markov-chain models sit at the high-interpretability, low-depth end of this space; deep-learning and graphical point-process models sit at the high-depth end, with lower or only moderate interpretability (Table 1). DBN-based latent-state models occupy a favorable interpretability position but, in their standard form, are constrained to first-order latent dynamics—placing them at high interpretability but limited temporal depth.

The gap addressed in this study is the absence, within the marketing-attribution literature, of a framework that retains the explicit, axiomatically transparent probabilistic structure of DBN-based latent-state models (Wei et al., 2023; Özyurt et al., 2022) while extending their temporal depth to represent the cumulative and delayed effects documented in customer journey research (Gao et al., 2021; Hatt & Feuerriegel, 2022) [6,7,10,14]. The three contributions introduced in Section I- (1) a DBN architecture linking touchpoints to latent engagement/intent states, (2) a non-Markovian, p-th-order extension of the latent-state dynamics, and (3) an interpretable counterfactual attribution mechanism operating on this extended latent structure—are designed to occupy this gap directly, positioning the proposed model at high interpretability and high temporal depth simultaneously.

### Research Hypotheses

**H1:** The proposed non-Markovian DBN provides lower attribution error than first-order Markov attribution models in customer journeys exhibiting cumulative exposure effects.

**H2:** The proposed framework captures delayed channel effects more accurately than conventional first-order DBN models.

**H3:** The proposed framework achieves competitive predictive performance while maintaining higher interpretability than deep-learning attribution approaches.

## 3. Proposed Methodology

### Customer Journey Representation

Each customer journey  $i$  is represented as a sequence of observed touchpoints,  $x_{i,1}, x_{i,2}, \dots, x_{i,T_i}$ , where each  $x_{i,t} \in \{1, \dots, M\}$  indicates which of  $M$  marketing channels (e.g., social media, search, email, display, mobile application) generated the  $t$ -th interaction, together with a terminal binary outcome  $y_i \in \{0, 1\}$  indicating whether the journey ended in conversion. We posit an unobserved sequence of latent behavioral states  $z_{i,1}, z_{i,2}, \dots, z_{i,T_i}$ , with  $z_{i,t} \in \{1, \dots, K\}$ , representing  $K$  latent engagement or intent states (for example, “unaware,” “considering,” and “high intent”). Each observed touchpoint  $x_{i,t}$  is assumed to be generated conditionally on the latent state  $z_{i,t}$ , and the conversion outcome  $y_i$  is assumed to depend on the latent-state trajectory, particularly its terminal value  $z_{i,T_i}$ .

### Dynamic Bayesian Network Specification

In a conventional first-order DBN, the model is fully specified by (i) an initial state distribution  $\pi(z_1)$ ; (ii) a transition model  $P(z_t | z_{t-1})$ , represented by a  $K \times K$  transition matrix  $A$ ; (iii) an emission model  $P(x_t | z_t)$ , represented by a  $K \times M$  emission matrix  $B$ ; and (iv) a conversion model  $P(y | z_{1:T})$  relating the latent trajectory to the conversion outcome. The joint probability of a journey factorizes as:

$$P(x_{1:T}, z_{1:T}, y) = \pi(z_1) \prod_{t=2}^T P(z_t | z_{t-1}) \prod_{t=1}^T P(x_t | z_t) P(y | z_{1:T}).$$

This factorization is the standard hidden Markov model / first-order DBN structure that underlies the latent-state customer behavior models.<sup>6,10</sup> Its key limitation for attribution purposes is that the transition model  $P(z_t | z_{t-1})$  embodies a one-step memory: any effect that an earlier touchpoint has on  $z_t$  must be mediated entirely through  $z_{t-1}$ , which cannot represent the kind of compounding, multi-step exposure effects (e.g., a second email reinforcing the effect of a first) that the marketing literature associates with cumulative campaign effects<sup>14</sup>.

### Non-Markovian Extension

The proposed model generalizes the transition component of the DBN to a  $p$ -th-order dependency on the latent-state

$$P(z_t' | z_{t-1}', z_{t-2}' \dots z_{t-p}').$$

When  $p = 1$ , this reduces to the conventional first-order DBN. For  $p > 1$ , the model can be represented computationally by augmenting the state space: defining a composite state

$$\tilde{z}_t = (z_t', z_{t-1}', \dots, z_{t-p+1}'),$$

which follows a first-order Markov chain over  $K^p$  possible composite values, allows standard DBN inference machinery to be applied to the augmented chain<sup>12</sup>. However, a full  $K^p \times K^p$  transition matrix is rarely identifiable or interpretable for  $K, p > 2$ . We therefore propose a parsimonious, autoregressive-style parameterization in which the transition probability is a function of a small number of lag-specific parameters—for example, a weighted combination of lag-specific transition matrices  $A_1, \dots, A_p$ , together with interaction terms capturing specific cumulative patterns (such as repeated exposure to the same channel across consecutive time steps). This preserves the substantive interpretability of the transition structure—each lag's contribution to the next latent state can be inspected directly—while keeping the number of free parameters manageable. Figure 1 (described narratively here, to be rendered graphically in the final manuscript) illustrates the structural difference: in the standard DBN, an arrow runs only from  $z_{t-1}$  to  $z_t$ ; in the proposed  $p$ -th-order extension, additional arrows run from  $z_{t-2}, \dots, z_{t-p}$  to  $z_t$ , representing the channels through which earlier touchpoints can exert delayed influence on the customer's current latent state.

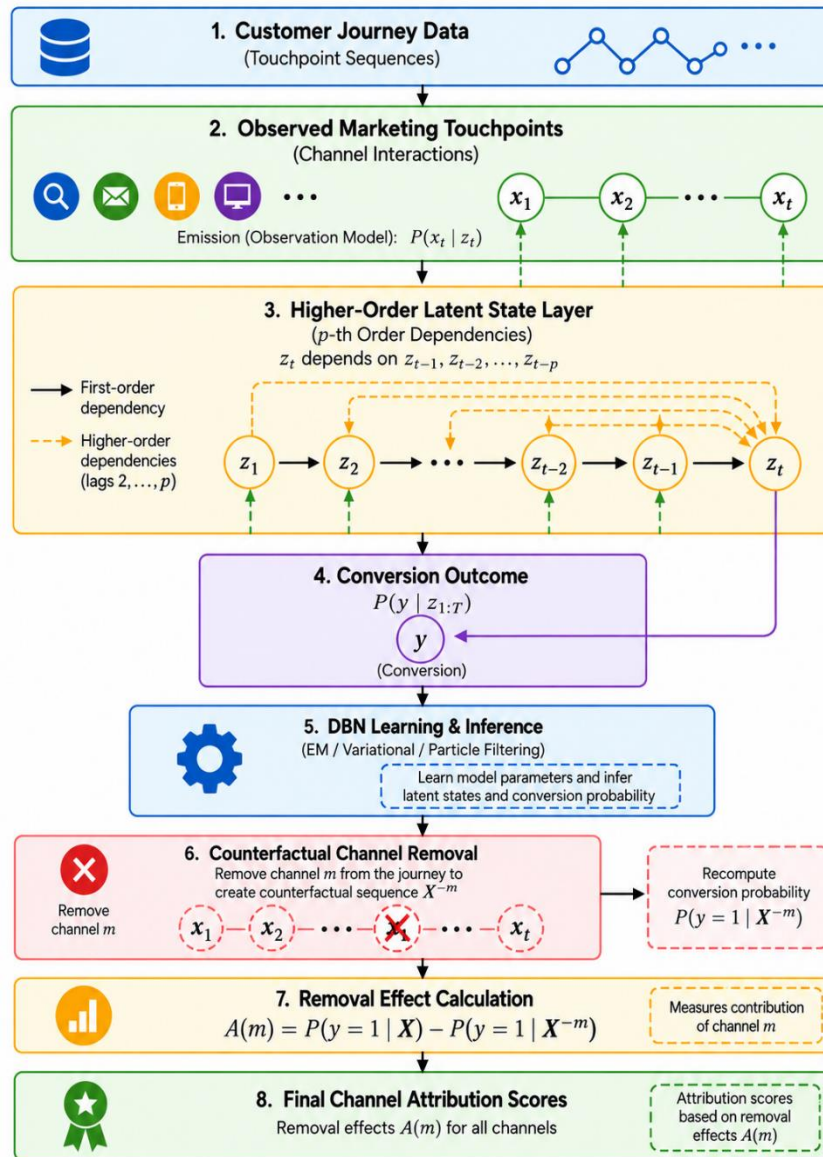


Figure 1. Proposed Non-Markovian Dynamic Bayesian Network Framework for Multi-Touch Attribution

## Inference and Learning

The parameter set  $\theta = \{\pi, A_1, \dots, A_p, B, \text{conversion-model parameters}\}$  is estimated from observed journey data via maximum-likelihood or maximum-a-posteriori estimation. For the augmented first-order representation, parameter learning can proceed via the Expectation-Maximization (EM) algorithm, using forward-backward (Baum-Welch-type) recursions over the composite state  $s_t$  during the E-step; Shiguihara et al.<sup>12</sup> provide a systematic overview of such exact and approximate DBN learning algorithms across application domains. Because the composite state space grows as  $K^p$ , exact forward-backward

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

recursions become computationally expensive for longer journeys, larger  $K$ , or higher orders  $p$ . We therefore propose approximate inference as the default for realistic deployments, via either (i) variational inference, which approximates the posterior over latent-state trajectories with a tractable factorized distribution, or (ii) sequential Monte Carlo / particle filtering, which represents the posterior with a set of weighted trajectory samples and updates them online as new touchpoints arrive. Chen and Li<sup>16</sup> review recent advances in differentiable particle filters that combine the flexibility of sampling-based inference with gradient-based parameter learning, making them a natural candidate for scalable estimation of the proposed non-Markovian latent dynamics. Model selection for the order  $p$  and the number of latent states  $K$  can be performed via held-out log-likelihood or information criteria, ideally combined with substantive input on the expected number and meaning of engagement states, in the spirit of the DBN elicitation<sup>13</sup>.

### Attribution via Counterfactual Removal-Effect Analysis

Channel-level attribution is defined as the change in expected conversion probability that results from a counterfactual intervention on channel  $m$ 's presence in the journey. Formally, for channel  $m$  we define the removal effect:

$$A(m) = P(y = 1 \mid x_{1:T}) - P(y = 1 \mid x_{1:T}^{(-m)}).$$

where  $x_{1:T}^{(-m)}$  denotes the journey with all occurrences of channel  $m$  replaced by a neutral “no-exposure” placeholder, and the expectation is taken over the fitted joint distribution of latent states and touchpoints. Because the latent-state transition model in the proposed framework depends on up to  $p$  preceding latent states—which are themselves shaped by the touchpoints observed at those times—removing channel  $m$  can alter not only the immediately following latent state but also latent states several steps later, including any cumulative-interaction terms (such as repeated-exposure effects) that involve channel  $m$ . This operationalizes, within an explicit probabilistic structure, the “delayed and cumulative effects” objective stated in Section I, and generalizes the removal-effect concept used in observed-state Markov chain attribution<sup>2</sup> and in the graphical point-process framework of Tao et al.<sup>5</sup> to operate on inferred behavioral states. Where axiomatic fairness properties across channel combinations are desired for budget-allocation purposes—for instance, to ensure that attributed credit sums exactly to the observed conversion rate—the channel-level removal effects  $A(m)$  can additionally be aggregated using a Shapley-value-consistent weighting scheme<sup>3</sup>.

## 4. Experimental Design

### Data Sources

Two complementary data sources are proposed. First, a real-world, anonymized multi-channel customer-journey dataset (e.g., from an e-commerce or retail setting) consisting of timestamped touchpoint sequences across the channels under study, together with a binary conversion label for each journey, will be used to evaluate predictive fit and the qualitative plausibility of estimated attribution patterns. Second, synthetic non-Markovian simulations—generated from a known  $p$ -th-order latent-state process with pre-specified, controllable channel effects—will be used to construct ground-truth attribution weights. As noted in Section I, real customer-journey data offers no ground truth for “correct” attribution; synthetic simulation is therefore essential for directly evaluating whether a model recovers known channel contributions, particularly contributions that arise only through higher-order, cumulative effects that a first-order model cannot, by construction, represent.

### Baseline Models

The proposed non-Markovian DBN will be compared against the following baselines, spanning the full range of approaches reviewed in Section II: (1) heuristic models—first-touch, last-touch, and linear attribution; (2) first-order Markov chain attribution with removal effects<sup>2</sup>; (3) higher-order Markov chain attribution, as a non-latent-variable alternative for capturing extended dependencies; (4) Shapley-value / game-theoretic attribution<sup>3</sup>; (5) a deep-learning sequence model in the spirit of the TCN-based architecture of Agrawal et al.<sup>4</sup>; (6) a standard first-order DBN / HMM with a single latent layer, representing the conventional latent-state baseline<sup>6,10</sup>; and (7), where feasible, the graphical point-process attribution framework of Tao et al. (2025)<sup>5</sup> as the closest recent non-Markovian alternative.

### Evaluation Metrics

Three families of metrics are proposed. First, attribution accuracy will be evaluated against the known channel-level contributions in the synthetic non-Markovian simulations, using the absolute and relative deviation between estimated and true removal effects. Second, classification-style metrics—precision, recall, and F1-score—will be computed for the task of identifying “high-impact” touchpoints or channels (defined, for example, as those whose true or estimated removal effect exceeds a threshold), allowing comparison across models on a common discriminative task even where exact attribution values differ in scale. Third, probabilistic goodness-of-fit will be assessed on the real-world dataset via held-out log-likelihood (or perplexity) of the conversion outcome and, where applicable, Kullback-Leibler divergence between model-implied and empirical distributions of journey outcomes. Computational cost (training and inference wall-clock time, and memory footprint as a function of journey length,  $K$ , and  $p$ ) will be reported as a secondary consideration informing the practical trade-offs summarized in Table 1.

## 5. Illustrative Simulation

To make the mechanism described in Section III concrete, this section presents a small, deliberately simplified numerical

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

illustration. It is intended purely to illustrate how a non-Markovian (second-order) latent-state model can attribute a larger removal effect to a channel with a cumulative, repeated-exposure effect than an otherwise identical first-order model; it is not an empirical result and does not draw on real customer data. All figures reported below were generated by Monte Carlo simulation (200,000 simulated journeys per condition) of a toy model with two latent states—Low intent (L) and High intent (H)—and three channels—Social, Search, and Email—over journeys of length  $T = 6$ .

In the toy model, the baseline probability of transitioning to the High-intent state is 0.30 from Low intent and 0.70 from High intent. Each touchpoint adds a channel-specific “boost” to this probability (Social: +0.02; Search: +0.05; Email: +0.15). In the second-order specification only, an additional cumulative boost of +0.20 is applied when both the current and immediately preceding touchpoints are Email and the latent state two steps earlier was already High—representing a delayed, compounding effect of repeated email exposure that a first-order model cannot represent by construction, since it has no access to  $z_{t-2}$ . Conversion probability is 0.80 conditional on ending in the High-intent state and 0.10 conditional on ending in the Low-intent state. The removal effect for email is computed as the difference in expected conversion probability between journeys drawn from the full channel-exposure distribution (Social 40%, Search 35%, Email 25%) and journeys in which Email exposures are replaced by neutral “no-exposure” touchpoints, with the remaining channels' relative shares preserved.

Table 2 reports the simulation results. Under the first-order specification, Email is attributed to the removal effect of 0.04207 (i.e., expected conversion probability falls from 0.52097 to 0.47890 when Email is removed). Under the second-order specification—identical in every respect except for the additional cumulative repeated-exposure term, Email's removal effect rises to 0.04863 (expected conversion probability falls from 0.52887 to 0.48024), an increase of approximately 15.6% relative to the first-order estimate ( $0.04863 / 0.04207 - 1 \approx 0.156$ ). This illustrates, in miniature, the central mechanism of the proposed framework: because the cumulative effect of repeated Email exposure operates through the dependence of  $z_t$  on  $z_{t-2}$  (via the second-order transition term), only a model that represents this dependency can attribute the corresponding share of conversion probability to Email. A first-order model, by construction, cannot represent this pathway and therefore systematically understates Email's true contribution in any data-generating process that includes such cumulative effects.

| Quantity                               | Order-1 model | Order-2 model | Difference        |
|----------------------------------------|---------------|---------------|-------------------|
| $E[P(\text{convert})]$ — Email present | 0.52097       | 0.52887       | +0.00790          |
| $E[P(\text{convert})]$ — Email removed | 0.47890       | 0.48024       | +0.00134          |
| Removal effect attributed to Email     | 0.04207       | 0.04863       | +0.00656 (+15.6%) |

Table no 2: Illustrative Simulation: Email Removal Effect Under First- and Second-Order Latent-State Models

Note. Figures are derived from a Monte Carlo simulation ( $N = 200,000$  journeys per condition,  $T = 6$  touchpoints) of a toy two-state, three-channel model described in the text. The simulation is illustrative only and does not represent results from real or full-scale synthetic experimental data as described in Section IV.

## 6. Expected Contributions and Positioning Relative to Baselines

Based on the theoretical mechanism described in Sections III and V and the empirical patterns documented in the reviewed literature, several expectations can be specified for the full empirical evaluation described in Section IV. First, on synthetic non-Markovian journeys, the proposed model is expected to recover channel-level attribution weights more accurately than first-order Markov chain, Shapley-value, and standard first-order DBN/HMM baselines, particularly for channels whose true contribution depends on repeated or delayed exposure—mirroring the mechanism illustrated in Section V and consistent with the documented benefits of duration-dependent and higher-order latent-state models over first-order HMMs in related clickstream and engagement-prediction<sup>6,7,11</sup>.

Second, relative to deep-learning attribution models such as the TCN-based architecture of Agrawal et al.<sup>4</sup>, the proposed model is expected to be competitive in predictive fit on real-world data while offering attribution values that are directly traceable to interpretable transition and emission probabilities, consistent with the interpretability-by-design properties associated with probabilistic graphical models in the taxonomy of Marcinkevičs and Vogt<sup>8</sup> and the broader call for trustworthy, decision-relevant explanations in Ali et al.<sup>9</sup>.

Third, relative to the graphical point-process framework of Tao et al.<sup>5</sup> the closest recent non-Markovian alternative—the proposed model is expected to offer a complementary rather than strictly superior representation: where the graphical point-process approach excels at representing relational excitation among many observed touchpoint types, the proposed DBN is expected to offer a more compact representation when the underlying driver of long-memory effects is plausibly a small number of latent engagement or intent states<sup>10</sup>, at potentially lower computational cost for journeys with a moderate number of channels. These expectations are framed as hypotheses to be tested in the full empirical evaluation; they are not presented as established findings.

## 7. Managerial Implications

If the expectations outlined in Section VI are borne out empirically, the proposed framework would offer marketing managers three practical capabilities beyond those provided by existing attribution tools. First, because attribution values are derived from explicit transition and emission probabilities associated with interpretable latent engagement states, managers can examine not only how much credit a channel receives, but through which latent-state pathway—for example, whether a channel's contribution operates primarily by moving customers from “unaware” to “considering,” or by reinforcing an already-high-intent state—supporting more targeted budget reallocation decisions than an aggregate attribution score alone<sup>1</sup>.

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

Second, the non-Markovian extension directly addresses a recurring practical concern in omnichannel marketing: that channels which build engagement cumulatively over multiple exposures (such as email nurture sequences) may be systematically undervalued by attribution models that consider only the most recent touchpoint or a single preceding state<sup>14</sup>. By explicitly representing how earlier exposures continue to influence the customer's latent state several steps later, the framework provides a basis for more accurately valuing such cumulative-effect channels in budget-allocation decisions.

Third, the latent-state representation offers a natural basis for designing customer engagement strategies that are conditioned on inferred intent state—for example, by identifying which channel-sequence patterns are most associated with transitions into high-intent states—supporting the kind of journey-level, rather than touchpoint-level, strategic planning advocated by Vollrath and Villegas<sup>1</sup> and Mahajan et al.<sup>15</sup>.

### 8. Limitations and Future Research

Several limitations should be acknowledged. First, the computational complexity of the proposed non-Markovian DBN increases with the number of latent states  $K$ , dependency order  $p$ , and journey length, potentially requiring approximate inference methods such as variational inference or particle filtering. Second, the interpretation and stability of unobserved latent states require careful validation and appropriate selection of  $K$  and  $p$ . Third, the framework has not yet been validated across diverse real-world industries, customer segments, or long customer journeys. Future research should therefore evaluate the model using larger real-world datasets, compare its attribution and predictive performance with established alternatives, and investigate scalable hybrid approaches that combine latent-state dynamics with graphical point-process representations.

### 9. Conclusion

This study proposes a non-Markovian Dynamic Bayesian Network (DBN) framework for multi-touch attribution that extends conventional first-order latent-state models to a  $p$ -th-order dependency structure. The key novelty is the integration of higher-order latent behavioral dynamics with counterfactual channel-level attribution, enabling earlier marketing exposures to influence subsequent latent states and conversion through delayed and cumulative pathways. Unlike conventional first-order Markov and DBN models, the proposed framework does not restrict the influence of historical exposures to the immediately preceding latent state.

Channel attribution is performed through counterfactual removal effects, allowing the contribution of individual channels to be evaluated through changes in expected conversion probability after removing their exposures. This provides an interpretable probabilistic mechanism for capturing cumulative effects while preserving the transparency of latent-state transition and emission structures. The illustrative simulation demonstrates the proposed mechanism. The Email removal effect increases from 0.04207 under the first-order model to 0.04863 under the second-order model, representing an approximately 15.6% increase. Although this experiment is illustrative rather than empirical, it demonstrates how higher-order latent dependencies can capture repeated-exposure effects that are not representable under a first-order formulation.

Overall, the proposed framework provides a novel balance between temporal depth, latent behavioral representation, and interpretability in multi-touch attribution. Its practical effectiveness, however, requires systematic validation using synthetic datasets with known attribution ground truth and real-world customer-journey data. Future research should evaluate attribution accuracy, predictive performance, computational scalability, and generalizability across industries and customer segments.

### References

1. Vollrath, M. D., & Villegas, S. G. (2022). Avoiding digital marketing analytics myopia: revisiting the customer decision journey as a strategic marketing framework. *Journal of Marketing Analytics*, 10(2), 106–113. <https://doi.org/10.1057/s41270-020-00098-0>
2. Gaur, J., Bharti, K., & Bajaj, R. (2026). Maximizing marketing impact: heuristic vs ensemble models for attribution modeling. *Global Knowledge, Memory and Communication*, 75(1–2), 359–377. <https://doi.org/10.1108/GKMC-04-2023-0112>
3. Molina, E., Tejada, J., & Weiss, T. (2022). Some game theoretic marketing attribution models. *Annals of Operations Research*, 318(2), 1043–1075. <https://doi.org/10.1007/s10479-022-04944-5>
4. Agrawal, A., Sheoran, N., Suman, S., & Sinha, G. (2022, April). Multi-touch attribution for complex B2B customer journeys using temporal convolutional networks. In *Companion Proceedings of the Web Conference 2022* (pp. 910–917). <https://doi.org/10.1145/3487553.3524670>
5. Tao, J., Chen, Q., Snyder Jr, J. W., Kumar, A. S., Meisami, A., & Xue, L. (2025). A graphical point process framework for understanding removal effects in multi-touch attribution. *Management Science*, 71(9), 7312–7332. <https://doi.org/10.1287/mnsc.2023.00457>
6. Özyurt, Y., Hatt, T., Zhang, C., & Feuerriegel, S. (2022, April). A deep Markov model for clickstream analytics in online shopping. In *Proceedings of the ACM Web Conference 2022* (pp. 3071–3081). <https://doi.org/10.1145/3485447.3512027>
7. Hatt, T., & Feuerriegel, S. (2022). Detecting user exits from online behavior: A duration-dependent latent state model. *arXiv preprint arXiv:2208.03937*. <https://doi.org/10.48550/arXiv.2208.03937>
8. Marcinkevičs, R., & Vogt, J. E. (2023). Interpretable and explainable machine learning: A methods-centric overview with concrete examples. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(3), e1493. <https://doi.org/10.1002/widm.1493>
9. Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., ... & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information Fusion*, 99, 101805. <https://doi.org/10.1016/j.inffus.2023.101805>
10. Wei, Q., Mu, Y., Guo, X., Jiang, W., & Chen, G. (2024). Dynamic Bayesian network-based product recommendation considering consumers' multistage shopping journeys: A marketing funnel perspective. *Information Systems Research*, 35(3), 1382–1402. <https://doi.org/10.1287/isre.2020.0277>
11. De La Torre, S. A., El Mistiri, M., Tung, K., Hekler, E., Klasnja, P., Pavel, M., ... & Marlin, B. (2025). A dynamic Bayesian network approach to modeling engagement and walking behavior: insights from a yearlong micro-randomized trial (Heartsteps II). *Health Psychology and Behavioral Medicine*, 13(1), 2552479. <https://doi.org/10.1080/21642850.2025.2552479>

## A Dynamic Bayesian Network Approach to Multi-Touch Attribution in Non-Markovian Customer Journeys

12. Shiguihara, P., Lopes, A. D. A., & Mauricio, D. (2021). Dynamic Bayesian network modeling, learning, and inference: A survey. *IEEE Access*, 9, 117639–117648. <https://doi.org/10.1109/ACCESS.2021.3105520>
13. Drury, K., & Smith, J. Q. (2024). Dynamic Bayesian Networks, Elicitation, and Data Embedding for Secure Environments. *Entropy*, 26(11), 985. <https://doi.org/10.3390/e26110985>
14. Gao, W., Fan, H., Li, W., & Wang, H. (2021). Crafting the customer experience in omnichannel contexts: The role of channel integration. *Journal of Business Research*, 126, 12–22. <https://doi.org/10.1016/j.jbusres.2020.12.056>
15. Mahajan, Y., Mahajan, V., & Kapse, M. (2024). Marketing analytics and consumer behavior: A systematic literature review for future research agenda. *Data-Driven Decision Making*, 151–167. [https://doi.org/10.1007/978-981-97-2902-9\\_7](https://doi.org/10.1007/978-981-97-2902-9_7)
16. Chen, X., & Li, Y. (2023). An overview of differentiable particle filters for data-adaptive sequential Bayesian inference. *arXiv preprint arXiv:2302.09639*. <https://doi.org/10.48550/arXiv.2302.09639>