informs
http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE

Submitted to *Marketing Science*

Vol. 00, No. 0, Xxxxx 0000, pp. 000–000

ISSN 0732-2399, EISSN 1526-548X

Modeling Behavioral Dynamics in Digital Content Consumption: An Attention-Based Neural Point Process Approach with Applications in Video Games

Junming Yin
Gotham AI, junming.yin.work@gmail.com

Yue (Katherine) Feng
Department of Management and Marketing, Faculty of Business, Hong Kong Polytechnic University, katherine.feng@polyu.edu.hk

Yong Liu
Department of Marketing, Eller College of Management, University of Arizona, yoliu@arizona.edu

Abstract. The consumption of digital content products (e.g., video games and live streaming) is often associated with multi-faceted, dynamically interacting consumer behavior that is subject to influence from pertinent external events. Inspired by these characteristics, we develop a novel attention-based neural point process approach to holistically capture the richness and complexity of consumer behavioral dynamics in modern digital content consumption. Our model features a new multi-representational, continuous-time attention mechanism that can flexibly model dynamic interactions between different types of behavior under external influence. Using learned representations as sufficient statistics of past events, we build a marked point process to efficiently characterize the occurrence time, behavior combination, and consumption quantity of consumers' future activities. We illustrate our model development and applications in the empirical context of a sports video game, showing its superior predictive performance over a wide range of baseline methods. Leveraging individual-level parameter estimates, we further demonstrate our model's utility for conducting segmentation analysis and evaluating the effects of past events on consumers' future engagement. Our model provides managers and practitioners with a powerful tool for developing more effective and targeted marketing strategies and gaining insights into consumer behavioral dynamics in digital content consumption.

Key words: digital content consumption; video game; behavioral dynamics; marked point process; attention mechanism; recurrent neural network

1. Introduction

The production and consumption of digital content is one of the most profound developments in business and technology in the last two decades. Fueled by advances in information technologies, products such as music streaming, on-demand movies and videos, live event streaming, and video games have become a prominent component of consumption by modern-day consumers. It is estimated that in 2020, the average consumer spent nearly 7 hours per day consuming digital content (DoubleVerify 2020), and the revenue of digital content industries exceeded $257 billion (Statista 2020). To capture business opportunities in these markets, it is imperative to develop new methods that can best harvest the value in the consumption data of increasing granularity and complexity to gain insights into consumer behavioral dynamics. Such insights can help content developers and marketing practitioners implement more effective product design and promotion strategies.

As compared with the consumption of traditional products and services, digital content consumption has several unique characteristics. First, consumer behavior in digital content consumption is often multi-faceted. A single session of content consumption activity often consists of various types of consumer behavior that may occur *concurrently*. For instance, consumers can leave live comments or tip video streamers when watching videos on a streaming platform, and video game players can make in-consumption purchases when playing a game session. The variety and concurrency of different behaviors greatly shape consumers' content consumption experience. Therefore, to understand consumer engagement, it is critical to model multiple types of consumption behavior jointly and their concurrent patterns, rather than examining each type of behavior separately. Second, digital content consumption often occurs episodically, with consumers accessing the content repeatedly over time, which results in complex intertemporal dynamics across different consumption activities. The multi-faceted nature of consumption behavior further complicates the dynamics. As a result, consumers' future engagement is no longer only determined by their past experiences in a single type of consumption behavior. A clear picture of consumer engagement should consider the dynamic interactions between the same and different types of consumption behavior. Third, consumer behavior in digital content consumption is often influenced by external events related to the product, particularly when the digital content has a real-world backdrop. For example, music artists' offline activities may affect consumers' music consumption in online streaming services, and players' video game consumption could be influenced by real-world events relevant to where the game is situated (e.g., sports, gardening, construction). To gain more accurate insights into consumer behavioral dynamics, it is crucial to account for the influence of pertinent external events.

Despite a growing body of research examining consumer behavioral dynamics in different digital content consumption contexts, no existing methods can be effectively applied to account for the aforementioned three characteristics. Most studies focus on modeling only one facet of consumption behavior and treat others as fixed variables without considering their dynamic interplay and potential external influences (e.g., Dew and Ansari 2018, Huang et al. 2019). To account for the variety of consumer behavior, one line of work in the business literature relies on classic multivariate point processes (Xu et al. 2014, Aggarwal et al. 2021), where each marginal process is calibrated to model one type of behavior. However, this class of statistical models (e.g., multivariate Hawkes processes (Hawkes 1971a,b)) typically makes strong assumptions, such as that interactions between different types of behavior are mutually exciting according to a particular parametric form, which may fail to capture the complexity of behavioral dynamics in digital content consumption. Moreover, the nature of multivariate point processes allows only one type of behavior to occur in a session of consumption activity, so they cannot directly model the concurrency of multiple behaviors.

In the more general context of modeling event dynamics (Aalen et al. 2008), an emerging stream of research in the machine learning literature has started to leverage the flexibility of recurrent neural network (RNN) to alleviate the potential model misspecification problem of classic point processes (e.g., Du et al. 2016, Mei and Eisner 2017). The main idea is to employ variants of RNN to learn a representation of the event history and use the learned representation as sufficient statistics of a marked point process to characterize the occurrence of future events. Although these approaches have improved performance over classic models in various event prediction tasks, they still suffer from several limitations when applied to model consumer behavioral dynamics in digital content consumption (see Section 5.1 for more details of these approaches). First, they heavily rely on RNN, which encodes a consumer's historical consumption information into a single hidden state vector to make future predictions. Consequently, they have limited capacity for capturing long-term influence from past events and complex dynamics of multiple behaviors. Second, when these approaches need to handle the concurrency of multiple behaviors, they must explicitly enumerate all possible behavior combinations (i.e., which subset of behaviors would occur in a consumption activity), leading to an exponential explosion in the number of parameters to be estimated. Third, in the presence of concurrent behaviors, the mark density function in these approaches needs to be used to specify the distribution of all possible behavior combinations. Hence, there is no systematic way to model and predict consumption quantity (e.g., the duration of song listening and video watching, purchase amount), which is highly indicative of a consumer's in-consumption engagement.[1]

In this paper, we develop a new attention-based neural point process approach to holistically account for the three essential characteristics of consumer behavior in digital content consumption. To capture the complex dynamics of multi-faceted consumption behavior under external influence, we first introduce a novel multi-representational, continuous-time attention mechanism on top of a widely used RNN architecture to learn multiple representations of consumption history.

[1] See Table EC.1 in Online Appendix A for a summary of more literature on modeling consumer behavioral dynamics.

Then we use the learned representations to build a univariate marked point process that couples the occurrence of multiple consumption behaviors. Moreover, we integrate hurdle models into the mark density function of the process to efficiently model the concurrency and consumption quantity of multiple behaviors. Our approach has three major innovations. First, unlike state-of-the-art RNN-based methods (e.g., Du et al. 2016, Mei and Eisner 2017) that encode a consumer's historical information into a single hidden state vector, our attention mechanism adaptively learns the relevance of all previous hidden states to form multiple representations of past events. Each representation identifies a unique and relevant aspect of the history for influencing each type of consumption behavior, thus significantly reducing irrelevant noises and increasing the flexibility to model dynamic interactions of multiple behaviors. Second, as consumption activities often occur at irregular time intervals, classic discrete-time attention mechanisms (e.g., Bahdanau et al. 2015) cannot account for continuous-time information in a consumer's consumption journey. We introduce a new time-decaying factor into the calculation of attention weights that can seamlessly incorporate rich elapsed time information into the learned representations of history. Third, the univariate nature of the process, together with custom-designed hurdle models, allows our method to efficiently handle the concurrency and all possible behavior combinations without an exponential explosion in the number of model parameters. The hurdle models also provide a systematic approach for characterizing and predicting the distribution of consumption quantity, thus generating a more comprehensive picture of consumer engagement.

We use a dataset from a major sports video game as an example of digital content consumption to illustrate the model development and applications. We focus on two main types of consumption behavior in video games, namely game-play and in-game purchase, and their dynamic interactions under the influence of relevant real-world professional sports matches. Based on players' past trajectories, our proposed approach can predict the occurrence time, behavior combination (i.e., login only, game-play only, purchase only, and concurrent game-play and purchase), and consumption quantity (i.e., game-play duration and purchase count) of future activities. Extensive experiments show that our model consistently outperforms a wide range of baseline methods, including two state-of-art recurrent marked point processes, a Bayesian survival model, a hidden Markov model, and other classic machine learning and statistical models. Various ablation studies are conducted to demonstrate the importance of incorporating each characteristic of digital content consumption into our framework. Leveraging individual-level parameter estimates, we further illustrate our model's utility of generating in-depth insights into consumer behavioral dynamics by performing consumer segmentation and evaluating the effects of past events. Our proposed approach provides a flexible and effective tool to help marketing practitioners and business managers better understand consumer behavior in digital content consumption and shed new light on their strategic designs and targeted marketing actions.

The remainder of the paper is organized as follows. We discuss the conceptual research background for our work in Section 2. Section 3 provides an overview of our dataset in the empirical context of sports video games. Model development, evaluation, and applications are described in Sections 4, 5, and 6. We conclude the paper with a discussion of managerial implications and avenues for future research in Section 7.

2. Research Background

In this section, we first summarize the essential characteristics of consumer behavior in digital content consumption that serve as the foundation and motivation for our model development. We then review the literature on player engagement, which provides further psychological and behavioral support for our concrete model specifications in the context of sports video games.

2.1. Digital Content Consumption

Digitization has revolutionized the production and consumption of media and entertainment products. The hedonic nature of these products, coupled with the fast and flexible delivery enabled by digital technologies, have made digital content consumption an on-demand and multi-tasking process. Consumers have unprecedented control over when, what, and how much to consume. Our discussion below focuses on three characteristics essential to capturing the richness and complexity of consumer behavior in digital content consumption: multi-faceted, dynamically interacting, and susceptible to external influence.

First, a typical business model of digital content consumption is characterized by various types of behavior that may occur concurrently, rather than just a single type of consumption behavior, in a session of consumption activity. For example, users can make live comments when watching online videos with a series of temporal variations in content (Zhang et al. 2020). The duration of video watching and the volume of live comments are highly indicative of a user's in-consumption engagement. Recent live-streaming platforms enable tipping behavior when viewers watch shows. Lu et al. (2021) examine the concurrency of tipping and viewing behaviors on a live-streaming platform and find that tipping behavior is influenced by the number of viewers. In the context of video games, as more features begin to be incorporated into games to improve the user experience, players are not restricted to only playing the core content in a game session. They may also engage in other types of behaviors, such as purchasing enhancement packs and communicating with other players.

Second, unlike many traditional products, such as consumer packaged goods, digital content does not disappear upon consumption. As compared to "finishing" a bottled drink, consumers can repeatedly access digital content (e.g., music, movies, video games) over time. Because of the hedonic and experiential nature of digital content consumption, a consumer's prior consumption experience can significantly influence subsequent engagement. The complexity of behavioral dynamics is further aggravated by the multifaceted nature of consumption behavior: different types of behaviors can dynamically interact with each other in an entangled way. For instance, Schweidel and Moe (2016) study the effects of previously viewed content and advertising on binge-watching behavior on a streaming video platform. The results indicate that viewers' previous viewing behavior stimulates future engagement, whereas viewers' response to advertisements (i.e., ads click) appears to discourage subsequent watching behavior. Dynamic interactions also exist between game-play and in-game purchase behaviors in the context of video games, as discussed in Section 2.2.

Third, due to the virtual nature of digital content, pertinent real-world events can affect digital content consumption. For instance, a music artist's social activities can influence how consumers listen to the artist's songs on streaming services. Martinez et al. (2021) show that the sociopolitical activism of popular singers has a significant impact on whether their songs would appear in a Spotify playlist. Similarly, the scores of real-world sports matches could affect players' behaviors (e.g., game-play and in-game purchase) in a sports video game. In the context of television advertising, Fossen and Schweidel (2019) find that product placement as a real-world promotional event can increase the volume of social media activities and website traffic for the featured brand. More generally, the interaction between online and offline events is prominent in omnichannel retail (Jing 2018, Tong et al. 2020, Ofek et al. 2011).

Our model simultaneously accounts for these important characteristics of digital content consumption. Though we illustrate its development and applications in the empirical context of sports video games, the modeling framework is very general and flexible. It can be easily adjusted or extended to other marketing contexts with multi-faceted, dynamically interacting consumption behavior under external influence.

2.2. Player Engagement in Video Games

Previous literature has identified various psychological factors that influence player engagement in video games. For example, applying self-determination theory (SDT; Deci and Ryan 1985), Ryan et al. (2006) find that the likelihood and duration of future game-play are associated with players' satisfaction of psychological needs for competence and relatedness in past in-game experiences. Another study by Colwell (2007) also confirms that multiple factors, including fun and stress relief, are major predictors of the frequency and duration of future game-play behaviors. In terms of players' repeat in-game purchase behaviors, the existing literature (e.g., Abdul-Muhmin 2010, Hsu et al. 2015) finds that factors such as satisfaction and trust are important drivers of consumers' future purchases. Therefore, players' game-play and purchase intentions are influenced by their past in-game experiences and can either grow or decay depending on the evolving direction of these factors. If previous game-play experiences decrease players' perceptions of their competence, or past in-game purchases cause frustration and dissatisfaction, players' motivation and future engagement may be undermined.

Moreover, researchers have documented the dynamic interaction between game-play and in-game purchase behaviors in the context of video games. For example, using an online survey, Mäntymäki and Salo (2011) demonstrate that recurring game-play can motivate players' future purchases of virtual goods. Analyzing online posts on two game bulletin boards, Lin and Sun (2011) observe mixed effects of in-game purchases on game-play activities: after making purchases, some players became more active, as they perceived improvements in game-play quality, whereas others decreased their participation, feeling that such purchases would weaken their appeal for fairness and put other players at a disadvantage. These findings suggest that game-play and in-game purchase behaviors can dynamically interact with variations in direction and strength. However, the existing literature on modeling the evolution of player engagement in video games considers only one type of behavior, either game-play (Huang et al. 2019) or purchase (Dew and Ansari 2018), while ignoring their dynamic and stochastic interactions.

Another group of factors that can shape player engagement relates to the influence of the real world. For example, Jung and Pawlowski (2014) interview 93 users from multiple virtual worlds, such as the online teenage community and online video games. Their study reveals that instead of completely isolating themselves in the virtual worlds, users consume virtual goods mainly to achieve real-world-oriented goals, such as seeking realistic experiences and simulating real-world activities. Specifically focusing on sports video games, Kim and Ross (2006) find that several factors, including knowledge and identification of real-world sports, are positively correlated with player engagement. Priming effects identified in social psychology provide further theoretical support for the influence of pertinent real-world events (Lashley 1951, Molden 2014). It has been shown that external stimuli can "prime" individuals to activate relevant concepts in their minds, with consequent effects on their subsequent behaviors (Wheeler et al. 2014). According to Yi (1990), such priming has an *affective component*, which triggers individuals' emotional reactions, and a *cognitive component*, which shapes their perceptions of relevant concepts (Srull and Wyer 1980, Erdley and D'Agostino 1988). In the context of video games, real-world matches from the professional sports league take the role of stimuli that may draw players' attention to related sports products. To account for potential influences from real-world sports matches, we adopt two important match characteristics from the league's official statistics, the number of highlights and the absolute score difference, to capture the affective and cognitive priming effects, respectively. The number of highlights indicates the number of exciting moments in a sports match[2], which can prime players' affective feelings. Previous research has found that excitement is a crucial factor for motivating player engagement (Kim and Ross 2006, Ryan et al. 2006). At the same time, the absolute score difference indicates the match result and can prime players' cognitive impressions of the strength of competing teams. As a result, players dissatisfied with a crushing loss in a recent match might be motivated to play virtual sports games to reduce their cognitive dissonance, while

[2] Each sports league has specific measures of exciting moments, such as three-pointers in basketball, shootings in soccer, and big plays in American football. We use the general term "the number of highlights" here to indicate the number of exciting moments in a match, which is usually released by the league's official statistics.

players whose expectations are confirmed by the result might also play virtual games to enhance their perceptions.

Motivated by the aforementioned findings, we develop a novel multi-representational attention-based neural point process to flexibly model players' complex behavioral dynamics under the influence of past in-game activities and external sports matches. Our model specification allows the effects of past events to be either stimulating or inhibiting, and their direction and strength can be inferred in a data-driven manner. The details of the model development are provided in Section 4.

3. Data Overview

We use a dataset from a leading American video game company[3] to illustrate the development of our model and its applications. The video game is designed and officially licensed to simulate the action and strategy of a real-world professional sports league. We observe two types of consumption behavior in the dataset: players can play game sessions in various modes (game-play) and may optionally purchase card packs to enhance their game experience (in-game purchase).

The dataset spans August 2014 to February 2015. Players with a minimum of five in-game activities were selected[4], resulting in a dataset of 2,020 players with a total of 143,509 in-game activities. For each in-game activity, we observe its occurrence time (i.e., login time), behavior combination (i.e., login only, game-play only, purchase only, and concurrent game-play and purchase), and consumption quantity (i.e., game-play duration and purchase count). The summary statistics of all in-game activities are reported in the upper part of Table 1. By examining the distribution of the behavior combination, we find that most in-game activities (75.57%) involve *game-play only*

[3] We are unable to reveal the name of the company and the video game because of a non-disclosure agreement.

[4] When evaluating the performance of all models (Section 5), we use 60%, 20%, and 20% of player activities in each event sequence for training, validation, and testing, respectively. Players with fewer than five in-game activities are not applicable in such an evaluation procedure, so they were removed. Only eight players with 27 in-game activities in total were not included in our study, accounting for less than 0.4% of all players.

Table 1 Event-Level Summary Statistics

Event	Characteristics/Marks	Mean	SD	Min	Max
In-game player activity	Game-play duration (in hours)	1.87	2.28	0	43.46
	Purchase count	0.44	1.21	0	30
Real-world sports match	Number of highlights	12.97	4.08	4	25
	Absolute score difference	12.61	9.54	0	52

Note. Number of in-game player activities $N = 143,509$. Number of real-world sports matches in the same time period $N_E = 268$.

without any enhancement pack purchase. In comparison, 2.88% are *purchase-only* activities, and 16.39% of in-game activities have *concurrent game-play and purchase*. The rest comprise *login-only* records (5.16%), probably because players were browsing the e-shop or visiting the video game community without any game-play or purchase.

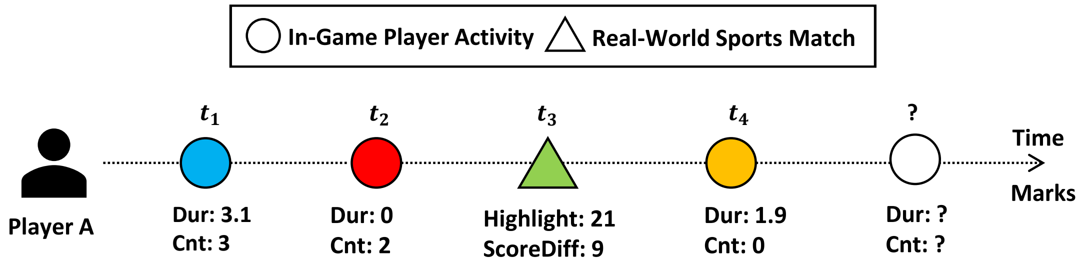
As discussed in Section 2.2, players' in-game activities may be influenced not only by their past activities but also by relevant real-world events. To investigate the influence of pertinent sports matches, we collected statistics on all 268 real-world matches from the professional sports league in the same time period. We consider the affective and cognitive priming effects from the external stimuli, and thus focus on the two important match characteristics available on the official league website: the number of highlights (i.e., exciting moments) and the absolute score difference between the competing teams. The summary statistics of these two variables are shown in the lower part of Table 1.

Table 2 Sequence-Level Summary Statistics

Variable	Mean	SD	Min	Max
Total number of in-game player activities	71.04	42.45	5	186
Total number of real-world sports matches	214.48	86.13	1	268
Total game-play duration (in hours)	133.20	116.28	0	901.42
Total purchase count	31.70	43.62	5	563

Note. Number of sequences/players $I = 2,020$.

Figure 1 **An Illustrative Example from Our Dataset of a Sports Video Game**



Note. For each player, we observe a sequence of in-game activities where the player can play game sessions or purchase card packs. For each in-game activity (depicted as circles), we observe the player's consumption quantity, i.e., game-play duration (Dur, in hours) and purchase count (Cnt). Real-world sports matches (depicted as triangles), with the number of highlights (Highlight) and absolute score difference (ScoreDiff) as the event characteristics, may influence players' in-game activities.

In our dataset, we have access to the timestamp when each player activated the game, which allows us to calculate the elapsed time of each event (i.e., in-game player activity or real-world sports match) since the activation of the game. Based on the occurrence time, we construct an event sequence for each player by sorting the events that the player has experienced in chronological order. Figure 1 provides an illustrative example of a player's event sequence with both in-game activities and real-world sports matches. The sequence-level summary statistics are provided in Table 2. On average, each event sequence includes about 71 in-game player activities and 215 sports matches.

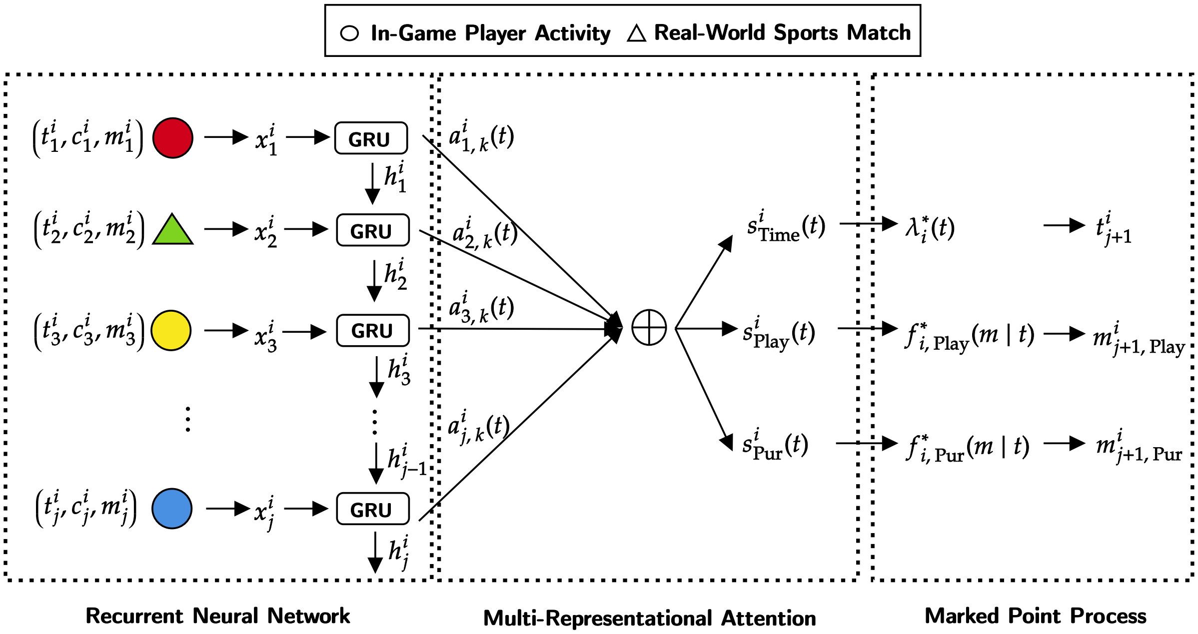
4. Model Development

We use the context of video games to illustrate our attention-based neural point process approach to modeling complex consumer behavioral dynamics under external influence. Our approach can be naturally extended to other digital content consumption contexts. In this section, we first provide a brief overview of our proposed model and then explain its different modules and estimation procedure in detail.

4.1. Overview

Consider we have observed the consumption journey of I video game players, where each player i has experienced N_i in-game activities and sports matches during our observation period. We can

Figure 2 **A High-Level Illustration of Our Proposed Model**



represent player i's event history with a sequence of vectors $\mathcal{S}_i = \{(t_j^i, c_j^i, \boldsymbol{m}_j^i)\}_{j=1}^{N_i}$ ordered by time, where t_j^i is the occurrence time of the j-th event, c_j^i denotes whether the event is an in-game activity $(c_j^i = 1)$ or a real-world sports match $(c_j^i = 2)$, and $\boldsymbol{m}_j^i$ is a vector (known as marks) capturing the consumption quantity of in-game activities or characteristics of real-world sports matches. For the major sports video game introduced in Section 3, the marks of an in-game activity include the game-play duration and purchase count, and the number of highlights and absolute score difference are used as marks of a real-world sports match (see Figure 1 and Table 1). Given player i's event history $\mathcal{H}_t^i = \{(t_j^i, c_j^i, \boldsymbol{m}_j^i) \mid t_j^i \leq t\}$, we aim to construct a probabilistic model to characterize the occurrence time and consumption quantity of future in-game activities.[5]

Figure 2 shows a high-level overview of our proposed model. The model consists of three key modules: a variant of RNN architecture known as GRU, multi-representational attention mecha-

[5] The occurrence time and marks of real-world sports matches are not stochastically modeled, as we assume that they are not influenced by virtual in-game activities. Therefore, we construct our probabilistic model of in-game activities conditioned on the observed real-world sports matches.

nism, and marked point process. GRU employs highly nonlinear mappings to encode a player's event history $\mathcal{H}_t^i = \{(t_j^i, c_j^i, \boldsymbol{m}_j^i) \mid t_j^i \le t\}$ into a sequence of hidden state vectors $\{\boldsymbol{h}_j^i\}$. Our novel multi-representational attention mechanism adaptively learns the relevance of $\{\boldsymbol{h}_j^i\}$ to form multiple continuous-time representations of past influence on future in-game activities: $\boldsymbol{s}_{\text{Time}}^i(t)$ for initiating an activity and $\{\boldsymbol{s}_{\text{Play}}^i(t), \boldsymbol{s}_{\text{Pur}}^i(t)\}$ for conducting each type of consumption behavior (i.e., game-play and purchase). The marked point process module uses the learned representations as sufficient statistics of the event history to parameterize its ground intensity and mark density functions. The ground intensity function $\lambda_i^*(t)$ characterizes the instantaneous occurrence rate of in-game activities and thus can be used to predict the next activity time t_{j+1}^i. The mark density function $\{f_{i,\text{Play}}^*(m \mid t), f_{i,\text{Pur}}^*(m \mid t)\}$ specifies the distribution of players' consumption quantity, i.e., game-play duration $m_{j+1,\text{Play}}^i$ and purchase count $m_{j+1,\text{Pur}}^i$.

4.2. Recurrent Neural Network

Our first step is to use a variant of RNN to extract nonlinear patterns from players' event history. For each event $(t_j^i, c_j^i, \boldsymbol{m}_j^i)$, we compute an embedding vector $\boldsymbol{x}_j^i$ as follows:

$$\boldsymbol{x}_j^i = \boldsymbol{W}_{c_j^i}[\Delta t_j^i, \boldsymbol{m}_j^i] + \boldsymbol{b}_{c_j^i}, \tag{1}$$

where $\Delta t_j^i = t_j^i - t_{j-1}^i$ is the elapsed time (with a week as the unit of time) since the last event.[6] Using two distinct sets of weight and bias parameters $\{\boldsymbol{W}_{c_j^i}, \boldsymbol{b}_{c_j^i}\}$ can account for the possibility that the influential strength of marks might be different for in-game activities ($c_j^i = 1$) and external events ($c_j^i = 2$). For example, an in-game activity with 5 hours of game-play and 3 purchases may have different influence on a player's subsequent behaviors compared with a sports match with 5 highlights and an absolute score difference of 3. Our estimation procedure described in Section 4.5 is able to infer the different weight and bias parameters in a data-driven manner.

[6] Recall from Section 3 that each player has a unique activation time in our dataset, so using the time interval between consecutive events can offset the different starting timestamps of each player's event sequence.

After this embedding step, the vectors $\{\boldsymbol{x}_j^i\}$ are of the same length for both in-game activities and sports matches, so we use them as the unified input to the RNN module. As a class of deep neural networks, RNN is designed to extract complex nonlinear patterns from sequential data into a set of hidden states. However, RNN is known to suffer from the "vanishing and exploding gradient problem" when handling long sequences (Goodfellow et al. 2016). Since many players in our dataset have long event sequences, we replace standard RNN with gated recurrent unit (GRU) (Cho et al. 2014), a widely used RNN variant, to avoid this issue. Compared with standard RNN, GRU incorporates two additional gating units, namely the update gate and reset gate, to selectively keep relevant information from the past and prevent the gradient from exploding/vanishing.[7] To apply GRU to our embedded event vectors $\{\boldsymbol{x}_j^i\}$, the hidden states $\{\boldsymbol{h}_j^i\}$ are recursively updated by processing the current input of the event and the previous hidden state:

$$\boldsymbol{h}_j^i = \mathrm{GRU}(\boldsymbol{h}_{j-1}^i, \boldsymbol{x}_j^i). \tag{2}$$

4.3. Multi-representational Attention Mechanism

Although GRU can alleviate the vanishing and exploding gradient problem, it encodes a player's historical information into a single hidden state vector, limiting its capacity for capturing long-term influence from past events and complex dynamics of multiple behaviors. Each hidden state of past events may have different relevance for future game-play and purchase, but GRU does not account for it. Moreover, when they are used to characterize the occurrence of future events, the discrete-time hidden states of GRU cannot incorporate elapsed continuous-time information in players' event history.

[7] The structure of GRU and its comparison with standard RNN are provided in Online Appendix B. We also develop a variant of our model with GRU replaced by the long short-term memory (LSTM, Hochreiter and Schmidhuber 1997). The two variants of our model show similar predictive performance, but the LSTM variant takes a much longer time and requires more memory to train because of its more complex structure. This is consistent with the conclusion from other sequence modeling tasks (Chung et al. 2014). Hence, we choose GRU as hidden units in our model.

To address these issues, we develop a novel multi-representational attention mechanism on top of GRU that adaptively learns the relevance of all hidden states to form multiple continuous-time representations of past influence on future in-game activities: $\boldsymbol{s}^i_{\text{Time}}(t)$ for initiating an activity, $\boldsymbol{s}^i_{\text{Play}}(t)$ for playing the game, and $\boldsymbol{s}^i_{\text{Pur}}(t)$ for making a purchase. Specifically, for each hidden state $\boldsymbol{h}^i_j$ in GRU, we use the following form to specify its influential scores $\{e^i_{j,\text{Time}}(t), e^i_{j,\text{Play}}(t), e^i_{j,\text{Pur}}(t)\}$ with respect to a future in-game activity at time t:

$$e^i_{j,k}(t) = \tanh\left(\boldsymbol{u}_k^\top \boldsymbol{h}^i_j \gamma_k(t - t^i_j)\right), \qquad k \in \{\text{Time, Play, Pur}\}, \tag{3}$$

where the term $\gamma_k(t) = e^{-w_k t}$ accounts for the time-decaying effect of past influence. By learning the parameters $\{\boldsymbol{u}_k\}$ and $\{w_k\}$, the influential scores $\{e^i_{j,\text{Time}}(t), e^i_{j,\text{Play}}(t), e^i_{j,\text{Pur}}(t)\}$ can adaptively determine how relevant is each previous hidden state $\boldsymbol{h}^i_j$ for initiating an in-game activity and conducting each type of behavior (i.e., game-play and purchase) at a future time t, with higher scores being more relevant. For each $k \in \{\text{Time, Play, Pur}\}$, our attention mechanism uses the normalized influential scores $\{a^i_{j,k}(t) \in (0,1)\}$, known as attention weights, to construct a weighted average of all previous hidden states $\{\boldsymbol{h}^i_j : t^i_j < t\}$:

$$\boldsymbol{s}^i_k(t) = \sum_{j:t^i_j<t} a^i_{j,k}(t)\boldsymbol{h}^i_j, \qquad \text{where} \qquad a^i_{j,k}(t) = \frac{\exp(e^i_{j,k}(t))}{\sum_{j:t^i_j<t} \exp\left(e^i_{j,k}(t)\right)}. \tag{4}$$

As a result, $\boldsymbol{s}^i_{\text{Time}}(t), \boldsymbol{s}^i_{\text{Play}}(t)$, and $\boldsymbol{s}^i_{\text{Pur}}(t)$ form continuous-time representations of past influence on initiating a future activity, playing the game, and making a purchase, respectively.

Compared with GRU, our proposed attention mechanism utilizes the information embedded in the hidden states $\{\boldsymbol{h}^i_j : t^i_j < t\}$ of all previous events to learn multiple representations of a player's event history. Each representation selectively identifies a unique and relevant aspect of the history for influencing each type of behavior, thus significantly reducing irrelevant noises and increasing the flexibility to capture dynamic interactions of multiple behaviors. Moreover, the time-decaying factor $\gamma_k(t)$ in Equation (3) incorporates rich elapsed time information $\{(t - t_j^i) \mid t_j^i < t\}$ from players' irregular consumption journey into the learned continuous-time representations, enabling them to be used as sufficient statistics of the history $\mathcal{H}_t^i \cup \{(t - t_j^i) \mid t_j^i < t\}$ for characterizing the occurrence of future in-game activities.

4.4. Marked Point Process

A marked point process is a continuous-time stochastic process that characterizes the occurrence rate of random events and the generative mechanism of associated marks.[8] Its multivariate extensions have been found successful in various applications to capture dynamic interactions between multiple behaviors (e.g., Xu et al. 2014, Aggarwal et al. 2021). However, their multivariate nature makes them unable to directly model concurrent game-play and purchase as well as login-only activities observed in our dataset (see Section 3). When each marginal process is used to model one type of behavior, players can only engage in either game-play or purchase when an in-game activity occurs. To address this issue, we use the learned representations of the event history to build a univariate marked point process that couples the occurrence of multiple behaviors. We also integrate hurdle models into the process to efficiently handle the concurrency and all possible behavior combinations without an exponential explosion in the number of model parameters.

Specifically, let $\mathcal{S}_i' = \{(t_j^i, c_j^i, \boldsymbol{m}_j^i) \mid c_j^i = 1\}_{j=1}^{N_i}$ be player i's in-game activity sequence. Each activity $(t_j^i, c_j^i, \boldsymbol{m}_j^i)$ can be viewed as a point with marks $\boldsymbol{m}_j^i$ occurring at t_j^i along the time line. Our goal is to learn the *ground intensity function* and *mark density function* of a marked point process to describe the generative mechanism of $\mathcal{S}_i'$ and future in-game activities. The ground intensity function $\lambda_i^*(t)$ characterizes the instantaneous occurrence rate of player i's in-game activities conditioned on the event history $\mathcal{H}_t^i = \{(t_j^i, c_j^i, \boldsymbol{m}_j^i) \mid t_j^i \leq t\}$ up to time t (including both in-game activities and real-world sports matches):

$$\lambda_i^*(t) \triangleq \lambda_i(t \mid \mathcal{H}_t^i) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(N_i'(t + \Delta t) - N_i'(t) > 0 \mid \mathcal{H}_t^i)}{\Delta t}, \tag{5}$$

[8] Online Appendix C provides a general introduction to marked point processes, including classic multivariate point processes.

where $N_i'(t)=\sum_{j=1}^{|\mathcal{S}_i'|}\mathbb{1}(t_j^i\leq t)$ is a counting process that counts the number of in-game activities up to time t. The mark density function $f_i^*(\boldsymbol{m}\mid t)\triangleq f_i(\boldsymbol{m}\mid\mathcal{H}_t^i,t)$ specifies the distribution of associated marks, namely game-play duration and purchase count.

Since the learned representations $\{\boldsymbol{s}_{\text{Time}}^i(t),\boldsymbol{s}_{\text{Play}}^i(t),\boldsymbol{s}_{\text{Pur}}^i(t)\}$ in Equation (4) encode the influence of past events on initiating a future activity, playing the game, and making in-game purchases, we can use them as sufficient statistics of the event history to parameterize the marked point process module. Specifically, we first use $\boldsymbol{s}_{\text{Time}}^i(t)$ to formulate the ground intensity function $\lambda_i^*(t)$:

$$\lambda_i^*(t)=\sigma_{\text{SP}}\left(\boldsymbol{v}_{\text{Time}}^\top\boldsymbol{s}_{\text{Time}}^i(t)+b_{\text{Time}}^i\right)=\sigma_{\text{SP}}\left(\sum_{j:t_j^i<t}a_{j,\text{Time}}^i(t)\boldsymbol{v}_{\text{Time}}^\top\boldsymbol{h}_j^i+b_{\text{Time}}^i\right),\tag{6}$$

where $\boldsymbol{v}_{\text{Time}}$ is a parameter to learn and b_{Time}^i specifies the baseline occurrence rate of player i's in-game activities to account for consumer heterogeneity. We choose the softplus specification $\sigma_{\text{SP}}(x)=\log(1+e^x)$ to guarantee that the ground intensity function is always positive. More importantly, our specification allows for both stimulating and inhibiting effects of past events: the sign of $\boldsymbol{v}_{\text{Time}}^\top\boldsymbol{h}_j^i$ determines whether event j increases or decreases the ground intensity, and thus the occurrence rate of in-game activities.

Next, we use $\{\boldsymbol{s}_{\text{Play}}^i(t),\boldsymbol{s}_{\text{Pur}}^i(t)\}$ as sufficient statistics to specify two hurdle models for the mark density function of game-play duration and purchase count. The hurdle models assume that players make decisions in a two-step manner. Each player i first decides whether to play or purchase with probability $\pi_{\text{Play}}^i(t)$ and $\pi_{\text{Pur}}^i(t)$, respectively, where

$$\pi_{\text{Play}}^i(t)=\sigma_{\text{S}}(\boldsymbol{v}_{\text{Play}}^\top\boldsymbol{s}_{\text{Play}}^i(t)+b_{\text{Play}}^i)\quad\text{and}\quad\pi_{\text{Pur}}^i(t)=\sigma_{\text{S}}(\boldsymbol{v}_{\text{Pur}}^\top\boldsymbol{s}_{\text{Pur}}^i(t)+b_{\text{Pur}}^i).\tag{7}$$

Here $\sigma_{\text{S}}(x)=e^x/(1+e^x)$ is the sigmoid function that ranges from 0 to 1. Given the decision to play the game, the game-play duration is then assumed to follow a Gamma distribution,

$$f_{i,\text{Play}}^*(m\mid t,m>0)=\frac{\beta^\alpha}{\Gamma(\alpha)}m^{\alpha-1}e^{-\beta m},\tag{8}$$

with α and β parameterized by $\boldsymbol{s}^i_{\text{Play}}(t)$, the past influence on future game-play,

$$\alpha \triangleq \alpha^i_{\text{Play}}(t) = \sigma_{\text{SP}}(\boldsymbol{v}^\top_{\text{Play},\alpha}\boldsymbol{s}^i_{\text{Play}}(t) + b^i_{\text{Play},\alpha}), \quad \beta \triangleq \beta^i_{\text{Play}}(t) = \sigma_{\text{SP}}(\boldsymbol{v}^\top_{\text{Play},\beta}\boldsymbol{s}^i_{\text{Play}}(t) + b^i_{\text{Play},\beta}). \tag{9}$$

Similarly, conditioned on the decision to make an in-game purchase, the purchase count follows a negative binomial distribution truncated at zero,

$$f^*_{i,\text{Pur}}(m \mid t, m > 0) = \frac{\Gamma(m+r)(1-p)^r p^m}{\Gamma(m)\Gamma(r)[1-(1-p)^r]}, \tag{10}$$

with r and p parameterized by $\boldsymbol{s}^i_{\text{Pur}}(t)$, the past influence on future purchase,

$$r \triangleq r^i_{\text{Pur}}(t) = \sigma_{\text{SP}}(\boldsymbol{v}^\top_{\text{Pur},r}\boldsymbol{s}^i_{\text{Pur}}(t) + b^i_{\text{Pur},r}), \qquad p \triangleq p^i_{\text{Pur}}(t) = \sigma_{\text{S}}(\boldsymbol{v}^\top_{\text{Pur},p}\boldsymbol{s}^i_{\text{Pur}}(t) + b^i_{\text{Pur},p}). \tag{11}$$

In summary, given behavior-specific representations $\{\boldsymbol{s}^i_{\text{Play}}(t), \boldsymbol{s}^i_{\text{Pur}}(t)\}$ of past influence, we specify the mark density function of game-play duration and purchase count as

$$f^*_{i,k}(m \mid t) = \begin{cases} 1 - \pi^i_k(t) & m = 0; \\ \pi^i_k(t) f^*_{i,k}(m \mid t, m > 0) & m > 0, \end{cases} \qquad k \in \{\text{Play, Pur}\}. \tag{12}$$

Such a specification of the mark density function, together with the univariate nature of our point process, allows us to efficiently model the concurrency and consumption quantity of multiple behaviors. Moreover, similar to the ground intensity function specified in Equation (6), the mark density function in Equations (7) to (12) also allows the effects of past events to vary in direction and strength: they could either increase or decrease the parameter values in the hurdle models (i.e., $\{\pi^i_{\text{Play}}(t), \alpha^i_{\text{Play}}(t), \beta^i_{\text{Play}}(t), \pi^i_{\text{Pur}}(t), r^i_{\text{Pur}}(t), p^i_{\text{Pur}}(t)\}$), and thus affect players' future game-play and purchase behaviors accordingly.

4.5. Parameter Estimation

The probabilistic characterization of the occurrence time and consumption quantity allows us to train our model with a maximum likelihood approach. Given the observed event sequences $\{\mathcal{S}_i\}_{i=1}^I$

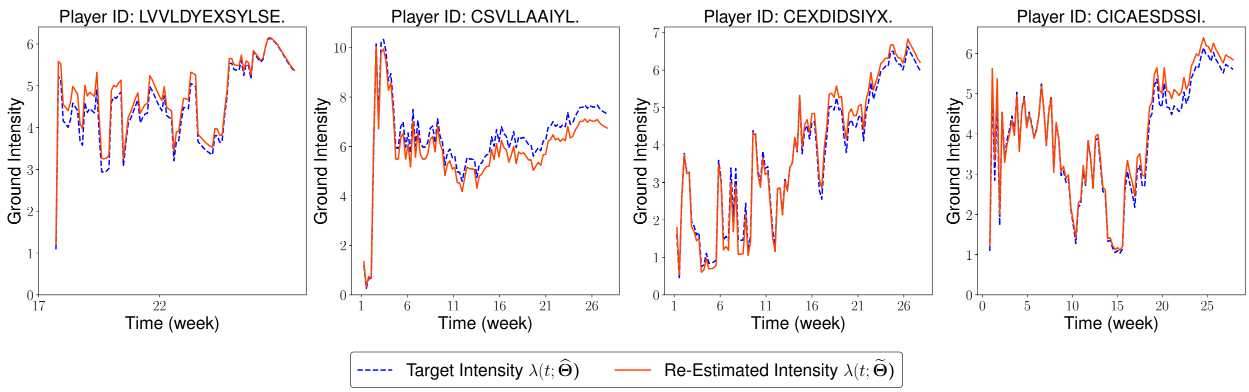
of I players, where $\mathcal{S}_i = \{(t^i_j, c^i_j, \boldsymbol{m}^i_j)\}_{j=1}^{N_i}$ includes inputs from both in-game activities ($c^i_j = 1$) and real-world sports matches ($c^i_j = 2$), the likelihood function can be written as (Daley and Vere-Jones 2003):

$$L(\boldsymbol{\Theta}) = \prod_{i=1}^{I} \left[\prod_{j:c^i_j=1}^{N_i} \lambda^*_i(t^i_j) \prod_{k\in\{\text{Play, Pur}\}} f^*_{i,k}(m^i_{j,k} \mid t^i_j) \right] \cdot \exp\left(-\int_0^T \lambda^*_i(\tau) d\tau \right), \tag{13}$$

where $\boldsymbol{\Theta}$ is a set of model parameters, the ground intensity function $\lambda^*_i(t)$ is given in Equation (6), and the mark density function $\{f^*_{i,\text{Play}}(m \mid t), f^*_{i,\text{Pur}}(m \mid t)\}$ is defined in Equation (12). It is worth noting that Equation (13) is the likelihood of in-game activities *only* because real-world sports matches are not stochastically modeled. However, real-world sports matches still contribute to the calculation of the likelihood function, as their occurrence time and marks are part of the event history used to specify the ground intensity and mark density of future in-game activities (see Equations (6) to (12)). As the stochastic integral $\int_0^T \lambda^*_i(\tau) d\tau$ does not have a closed-form expression with our specification of the ground intensity function, we apply the Monte Carlo method to evaluate such term. The algorithmic details are provided in Online Appendix D. We further employ the Adam optimizer (Kingma and Ba 2015) to maximize the likelihood and estimate the parameters $\boldsymbol{\Theta}$.

We conduct the following simulation studies to demonstrate the accuracy of our estimation procedure. First, we use the original sports video game dataset to obtain the estimated model parameters $\widehat{\boldsymbol{\Theta}}$. Next, we extend Ogata's thinning algorithm (Ogata 1981) and develop a simulation algorithm to simulate $I = 2,020$ sequences of in-game activities (the same sample size as the original dataset) from our model with the parameters $\widehat{\boldsymbol{\Theta}}$. The details of the simulation algorithm are provided in Online Appendix E. We then re-estimate the model parameters, denoted as $\widetilde{\boldsymbol{\Theta}}$, based on the simulated sequences. If our estimation procedure is accurate, the re-estimated intensities $\lambda(t; \widetilde{\boldsymbol{\Theta}})$ should closely approximate the target ground intensities $\lambda(t; \widehat{\boldsymbol{\Theta}})$. Figure 3 shows a visual

Figure 3 Comparison of the Target and Re-Estimated Ground Intensity Functions



Note. Blue dotted lines: the target ground intensity functions $\lambda(t;\widehat{\mathbf{\Theta}})$ are evaluated with the model parameters $\widehat{\mathbf{\Theta}}$ estimated from the original sports video game dataset. Red solid lines: the re-estimated ground intensity functions $\lambda(t;\widetilde{\mathbf{\Theta}})$ are evaluated with the model parameters $\widetilde{\mathbf{\Theta}}$ estimated from the simulated datasets. The simulated datasets are generated from our point process with the model parameters $\widehat{\mathbf{\Theta}}$, using the simulation algorithm in Online Appendix E.

comparison of the target and re-estimated intensity functions for four randomly selected players in a representative run of the simulation, demonstrating that our estimation procedure can successfully recover target ground intensity functions of different trends. We present a more qualitative comparison and quantitative evaluation based on residual analysis in Online Appendix F.

5. Model Evaluation

Our proposed framework, based on an attention-based neural point process, provides a principled way to *simultaneously* forecast the dynamics of multiple behaviors in digital content consumption. In this section, we first demonstrate its superior performance over a wide range of baselines in predicting the occurrence time, behavior combination, and consumption quantity (i.e., game-play duration and purchase count) of future in-game activities on our sports video game dataset. We then conduct ablation studies to illustrate the importance of introducing the multi-representational attention mechanism, modeling the dynamic interplay between game-play and purchase behaviors, and accounting for the influence from external sports matches.

When evaluating the predictive performance of each model, we split the sports video game dataset introduced in Section 3 into training, validation, and test sets in chronological order so that they include 60%, 20%, and 20% of total in-game activities, respectively.[9] We first train each model on the training set with different hyperparameter values. Then we select the optimal values of hyperparameters that achieve the best predictive performance on the validation set. For RNN-based models, the hyperparameters to be tuned include the size of the embedding layer {32, 64, 128, 256} (i.e., the dimension of $\boldsymbol{x}_j^i$), the size of hidden layer {32, 64, 128, 256} (i.e., the dimension of $\boldsymbol{h}_j^i$), and the initial learning rate for the Adam optimizer {0.00001, 0.001, 0.01}. Finally, each model trained with the optimal values of hyperparameters is evaluated based on its out-of-sample prediction on the test set.

5.1. Predicting Future In-Game Activities

5.1.1. Predicting the Occurrence Time. As the ground intensity function $\lambda_i^*(t)$ in Equation (6) characterizes the instantaneous occurrence rate of player i's in-game activities, we use it to predict the occurrence time $\hat{t}_{j+1}^i$ of the player's $(j+1)$-th activity based on the conditional expectation:

$$\hat{t}_{j+1}^i = \mathbb{E}\left[t_{j+1}^i \middle| \mathcal{H}_{t_j^i}^i\right] = \int_{t_j^i}^{\infty} \tau \cdot f_i^*(\tau) d\tau. \tag{14}$$

Here, $f_i^*(t)$ is the probability density function of the next activity time conditioned on the player's event history $\mathcal{H}_{t_j^i}^i$, which is related to the ground intensity function $\lambda_i^*(t)$ as $f_i^*(t) = \lambda_i^*(t) \exp\left(-\int_{t_j^i}^t \lambda_i^*(\tau) d\tau\right)$ (Daley and Vere-Jones 2003). We compare the predictive performance of our model against two state-of-the-art recurrent marked point processes, two classic proportional hazard models, and a Bayesian survival model.

- Recurrent Marked Temporal Point Processes (RMTPP) (Du et al. 2016): RMTPP is one of the first recurrent marked point processes to refrain from the strong assumptions made in classic models, such as Hawkes processes. RMTPP employs a standard RNN to learn a discrete-time

[9] We also split the dataset by individuals to demonstrate the performance of our model in predicting the behaviors of out-of-sample players. The results are provided in Online Appendix G.

representation of the event history and uses it to parameterize the ground intensity function of a point process. In our experiment, we replace the standard RNN with GRU for a fair comparison.

- Neural Hawkes (Mei and Eisner 2017): Neural Hawkes is another type of recurrent marked point process. In contrast to RMTPP, Neural Hawkes proposes a continuous-time LSTM to learn a time-varying representation of the event history. To apply it to our context, we use the learned representation to parameterize four marginal point processes, one for each distinct behavior combination of in-game activities (i.e., login only, game-play only, purchase only, and concurrent game-play and purchase). The total ground intensity of the process is given by the sum of four marginal intensity functions.
- Proportional Hazard Models (PHMs) (Seetharaman and Chintagunta 2003): PHMs are classic continuous-time stochastic models that have been applied to investigate how covariates summarized from the event history affect the occurrence time of individual behaviors. The ground intensity function of PHMs is specified as $\lambda_i^*(t) = \lambda_0(t)e^{X_t^i\beta}$, where the baseline hazard $\lambda_0(t)$ is a univariate function of time and the covariate vector X_t^i contains the summary statistics of the event history before time t.[10] For the baseline hazard $\lambda_0(t)$, we choose two parametric specifications for comparison: Weibull PHM (WB PHM), where $\lambda_0(t) = \gamma\alpha(\gamma t)^{\alpha-1}$; Log-logistic PHM (LL PHM), where $\lambda_0(t) = \gamma\alpha(\gamma t)^{\alpha-1}/(1+(\gamma t)^{\alpha})$.
- Bayesian Survival Model (BSM):[11] We consider a Bayesian survival model in which the cumulative distribution function (CDF) of player i's next activity time conditioned on the event history $\mathcal{H}^i_{t^i_j}$, i.e., $F_i^*(t) \triangleq F(t \mid \mathcal{H}^i_{t^i_j}) = \int_{t^i_j}^t f_i^*(\tau)d\tau$, has a Beta-Stacy process prior (Walker and Muliere 1997). That is, $F_i^*(t) \sim \mathcal{BS}(c_i^*(t), G_i^*(t))$, where the scale function $c_i^*(t)$ controls the dispersion

[10] Specifically, we follow Fader et al. (2005) to create a combination of recency (the occurrence time of the most recent in-game activity and sports match), frequency (the number of prior in-game activities and sports matches), and monetary value (average mark values per event) as the covariate vector. The same covariate vector is used for all the non-RNN-based baseline models.

[11] We thank an anonymous reviewer for suggesting this baseline.

Table 3 Performance of Predicting the Occurrence Time of In-Game Activities

Model	RMSE	MAE
Our Model	**0.42**	**0.35**
RMTPP	0.92	0.38
Neural Hawkes	0.69	0.53
WB PHM	2.63	0.71
LL PHM	2.64	0.67
BSM	1.25	0.62

of the random CDFs around the mean function $G_i^*(t)$. In our specification, we follow the common practice in the literature (Arfè et al. 2021) to assume $c_i^*(t) = 1$. We further choose a Weibull CDF (Rigat and Muliere 2012) for the mean function $G_i^*(t) = 1 - e^{-X_t^i \beta}$, which is parameterized with the same summary statistics X_t^i of the event history used in PHMs.

Table 3 reports the results of predicting the occurrence time of in-game activities. Our proposed model achieves the lowest root mean squared error (RMSE) and mean absolute error (MAE). The Bayesian survival model performs much better than classic proportional hazard models, i.e., Weibull PHM and Log-logistic PHM, because of its greater flexibility in modeling the occurrence time of in-game activities. However, there is still a considerable performance gap between the Bayesian survival model and RNN-based approaches (our model, RMTPP, and Neural Hawkes), which is mainly due to the greater capacity of RNN to extract more useful information from players' event history. Compared with RMTPP and Neural Hawkes that heavily rely on RNN, the novel continuous-time attention mechanism employed on top of RNN in our model is able to learn a more effective representation of a player's entire event history (i.e., $\boldsymbol{s}^i_{\text{Time}}(t)$) that is relevant for making predictions of future time. Consequently, our model significantly outperforms these two baselines.

5.1.2. Predicting the Behavior Combination. After obtaining the estimated occurrence time $\hat{t}^i_{j+1}$ of the $(j+1)$-th in-game activity, we then predict whether player i plays or purchases in the activity. We use $\boldsymbol{\omega}^i_{j+1} = \left(\omega^i_{j+1,\text{Play}}, \omega^i_{j+1,\text{Pur}}\right)$ to represent the behavior combination of the $(j+1)$-th in-game activity, with $\omega^i_{j+1,\text{Play}} = 1$ indicating that player i would play and $\omega^i_{j+1,\text{Pur}} = 1$ indicating that player i would purchase. This leads to four possible combinations of game-play and purchase behaviors (i.e., login only, game-play only, purchase only, and concurrent game-play and purchase). As $\pi^i_{\text{Play}}(t)$ and $\pi^i_{\text{Pur}}(t)$ in Equation (7) specify the probability of player i to play and purchase at time t, we combine them with the estimated occurrence time $\hat{t}^i_{j+1}$ to predict the behavior combination $\hat{\boldsymbol{\omega}}^i_{j+1} = \left(\hat{\omega}^i_{j+1,\text{Play}}, \hat{\omega}^i_{j+1,\text{Pur}}\right)$ of the next in-game activity:

$$\hat{\omega}^i_{j+1,\text{Play}} = \mathbb{1}\left[\pi^i_{\text{Play}}(\hat{t}^i_{j+1}) > 0.5\right] \quad \text{and} \quad \hat{\omega}^i_{j+1,\text{Pur}} = \mathbb{1}\left[\pi^i_{\text{Pur}}(\hat{t}^i_{j+1}) > 0.5\right]. \tag{15}$$

We compare the performance of our model against a set of relevant baselines, including RMTPP, Neural Hawkes, Hidden Markov Model, Logistic Regression, and Naïve Bayes:

- RMTPP (Du et al. 2016): Using the learned discrete-time representation as sufficient statistics of the event history, RMTPP employs a softmax layer (i.e., a multinomial distribution) to output the probability of each possible behavior combination. Because of the discrete-time nature of the learned representation, the predicted probability does not depend on the estimated occurrence time $\hat{t}^i_{j+1}$ of the next activity, so it cannot take into account the elapsed continuous-time information from the most recent event in the history $\mathcal{H}^i_{t^i_j}$ (i.e., $\hat{t}^i_{j+1} - t^i_j$).
- Neural Hawkes (Mei and Eisner 2017): Based on the learned continuous-time representation of the event history, Neural Hawkes parameterizes four marginal point processes, one for each possible behavior combination. The predicted probability of each behavior combination is proportional to its corresponding marginal intensity evaluated at the estimated occurrence time $\hat{t}^i_{j+1}$ of the next activity.
- Hidden Markov Model (HMM) (Rabiner 1989):[12] Following Huang et al. (2019), we specify a discrete-time HMM in which the discrete latent state variables model the evolution of players'

[12] We thank an anonymous reviewer for suggesting this baseline.

Table 4 Performance of Predicting the Behavior Combination of In-Game Activities

Model	Accuracy	Micro-F1	Macro-F1
Our Model	**0.76**	**0.88**	**0.62**
RMTPP	0.74	0.87	0.49
Neural Hawkes	0.73	0.86	0.53
HMM	0.70	0.75	0.50
LR	0.67	0.68	0.48
NB	0.67	0.66	0.55

engagement level. To capture players' game-play and purchase behaviors in different engagement levels, we use two latent-state-dependent hurdle models (Gamma and negative binomial truncated at zero) to model the distribution of game-play duration and purchase count, respectively. We apply the same decision rule in Equation (15) to predict the behavior combination of the next in-game activity.

- Logistic Regression (LR) and Naïve Bayes (NB): We also train two classic machine learning models to predict the behavior combination, with the same summary statistics X_t^i of the event history used in PHMs as the feature vector.

Table 4 presents the accuracy and micro/macro-averaged F1-score of all competing methods in predicting the behavior combination. The macro-averaged score is an arithmetic mean of F1-scores for the prediction of whether an in-game activity involves game-play or purchase. By contrast, the micro-averaged score weights F1-scores by the number of in-game activities with game-play or purchase, respectively. The results show that our model achieves the best performance in all three metrics, with the largest improvement (7% increase) on the Macro-F1 score over the second-best method. This is because other baselines poorly predict the occurrence of less frequent purchase behaviors, and the Macro-F1 score assigns equal importance to the correct prediction of future game-play and purchases. As compared with RMTPP and Neural Hawkes, the effective specification of hurdle models based on the learned behavior-specific representations of past influence (i.e., $\{\boldsymbol{s}^i_{\text{Play}}(t), \boldsymbol{s}^i_{\text{Pur}}(t)\}$) allows our model to predict the behavior combination more accurately without an exponential explosion in the number of parameters.

5.1.3. Predicting the Consumption Quantity. Finally, given the predicted behavior combination $\hat{\boldsymbol{\omega}}^i_{j+1}$, we forecast the consumption quantity of the next in-game activity, i.e., the game-play duration $\hat{m}^i_{j+1,\text{Play}}$ and purchase count $\hat{m}^i_{j+1,\text{Pur}}$. Specifically, for $k \in \{\text{Play, Pur}\}$, we predict $\hat{m}^i_{j+1,k} = 0$ if $\hat{\omega}^i_{j+1,k} = 0$ and use the following conditional expectations if $\hat{\omega}^i_{j+1,k} = 1$, which have closed-form expressions with our specified mark density function in Equations (8) and (10):

$$\hat{m}^i_{j+1,\text{Play}} = \mathbb{E}\left[m^i_{j+1,\text{Play}} \middle| \hat{t}^i_{j+1}, \hat{\omega}^i_{j+1,\text{Play}} = 1, \mathcal{H}^i_{t^i_j}\right] = \frac{\alpha^i_{\text{Play}}(\hat{t}^i_{j+1})}{\beta^i_{\text{Play}}(\hat{t}^i_{j+1})}, \tag{16}$$

$$\hat{m}^i_{j+1,\text{Pur}} = \mathbb{E}\left[m^i_{j+1,\text{Pur}} \middle| \hat{t}^i_{j+1}, \hat{\omega}^i_{j+1,\text{Pur}} = 1, \mathcal{H}^i_{t^i_j}\right] = \frac{p^i_{\text{Pur}}(\hat{t}^i_{j+1}) r^i_{\text{Pur}}(\hat{t}^i_{j+1})}{\left(1 - p^i_{\text{Pur}}(\hat{t}^i_{j+1})\right)\left(1 - (1 - p^i_{\text{Pur}}(\hat{t}^i_{j+1}))^{r^i_{\text{Pur}}(\hat{t}^i_{j+1})}\right)}.$$

We compare our method with the classic hurdle models and HMM for model evaluation.

- Hurdle Models (Cameron and Trivedi 2013): We use the Gamma hurdle model (Gamma Hurdle) and negative binomial hurdle model (NB Hurdle) to predict future game-play duration and purchase count, respectively. The hurdle models have a similar probabilistic specification as in Equations (7) to (12), except that the sufficient statistics vector $\{\boldsymbol{s}^i_{\text{Play}}(t), \boldsymbol{s}^i_{\text{Pur}}(t)\}$ is replaced with the covariate vector X^i_t of the event history (up to time t) used in PHMs.
- Hidden Markov Model (HMM) (Rabiner 1989): As HMM has two latent-state-dependent hurdle components, it can also be used to predict future consumption quantity.

The results in Table 5 clearly demonstrate the advantage of our approach in flexibly modeling the complex dynamics and efficiently learning rich information from the event history. Although all competing methods share the same probabilistic distributions for game-play duration and purchase count, they use historical information differently. Hurdle models use a simple summary statistics vector of the event history; HMM incorporates such a vector to capture the evolution of discrete latent state (i.e., players' engagement level); our model learns multiple sufficient statistics of past events, creating a unique representation of the event history for influencing each type of consumption behavior (i.e., $\boldsymbol{s}^i_{\text{Play}}(t)$ for game-play and $\boldsymbol{s}^i_{\text{Pur}}(t)$ for purchase). Moreover, as compared with

Table 5 Performance of Predicting the Consumption Quantity of In-Game Activities

Model	Game-Play Duration		Purchase Count	
	RMSE	MAE	RMSE	MAE
Our Model	**2.05**	**1.30**	**1.25**	**0.45**
Gamma Hurdle	5.55	1.83	-	-
NB Hurdle	-	-	5.48	0.71
HMM	2.45	1.57	1.48	0.66

HMM, our attention-based approach has a much higher model capacity: (1) the hidden states in our model are continuous high-dimensional vectors, which can store much richer information of the dynamics than a few discrete hidden states in HMM; (2) HMM uses simple transition rules, whereas our model employs highly nonlinear mappings to update hidden states, allowing it to more flexibly adapt to the complex behavioral dynamics of in-game activities.

5.2. Ablation Studies

We developed our novel framework to model complex player behavior that is multi-faceted, dynamically interacting, and subject to external influence. To demonstrate the importance of various design choices in our approach that can jointly capture these essential characteristics, we conduct the following ablation studies.

First, our model introduces a new multi-representational attention mechanism to adaptively learn a unique representation of past events for influencing each type of consumption behavior, e.g., $\boldsymbol{s}^i_{\text{Play}}(t)$ for game-play and $\boldsymbol{s}^i_{\text{Pur}}(t)$ for purchase. As a result, the model allows the dynamics of each behavior to be influenced by a distinct and relevant aspect of the event history. To verify the effectiveness of our multi-representational attention mechanism, we replace it with a standard single-representational attention mechanism, such that only a single sufficient statistics $\boldsymbol{s}^i(t) = \sum_{j:t^i_j<t} a^i_j(t)\boldsymbol{h}^i_j$ of the event history is learned to parameterize the dynamics of both game-play and purchase behaviors. The predictive performance of the resulting model drops, as shown in Table 6, demonstrating the importance of multi-representational learning in modeling multi-faceted consumption behavior.

Table 6 Results of the Ablation Studies

Setting	Occurrence Time		Behavior Combination		Game-Play Duration		Purchase Count	
	RMSE	MAE	Micro-F1	Macro-F1	RMSE	MAE	RMSE	MAE
(0)	**0.42**	**0.35**	**0.88**	**0.62**	**2.05**	**1.30**	**1.25**	**0.45**
(1)	0.65	0.37	0.87	0.55	2.09	1.33	1.29	0.46
(2.1)	0.54	0.42	-	-	2.12	1.44	-	-
(2.2)	0.62	0.48	-	-	-	-	1.46	0.52
(3)	0.60	0.41	0.87	0.53	2.16	1.39	1.35	0.48

Note. (0): full model and full dataset; (1): alternative model with single-representational attention mechanism and full dataset; (2.1): alternative model that treats purchase count as fixed and full dataset; (2.2): alternative model that treats game-play duration as fixed and full dataset; (3): full model and partial dataset without sports matches.

Second, a unique advantage of our method is that it can flexibly model the dynamic interplay between game-play and purchase behaviors. To illustrate its effectiveness, we conduct another ablation study by training two variants of our model on the full dataset. Specifically, each model variant "turns off" the stochastic nature of one behavior, either game-play or purchase, and treats its marks as fixed inputs. The resulting model variants still use the same ground intensity function as in Equation (6), but consider only the mark density function of one behavior (see Equation (12)) in the likelihood. In this way, one type of behavior cannot be influenced by and predicted from past events, so the model can only account for the unidirectional influence between game-play and purchase behaviors. The results in Table 6 clearly show that effectively capturing dynamic interactions can lead to a more accurate prediction of future in-game activities.

Third, our model includes real-world sports matches as part of players' event history to learn the ground intensity and mark density functions of in-game activities (see Equations (6) to (12)). To show the benefit of incorporating related external influence, we construct another event sequence $\mathcal{S}_i' = \{(t_j^i, c_j^i, m_j^i)\}_{j=1}^{N_i'}$ for each player i from the original dataset, with all sports matches removed (i.e., $c_j^i = 1$ for all the events). We then retrain our model based on this partial dataset that contains the sequences of in-game activities only. The performance gap between the original model and the new retrained model in Table 6 provides further evidence that the influence from related real-world events is one of the major factors for shaping player engagement in video games, as discussed in Section 2.2.

6. Model Applications

Our model shows superior performance in predicting players' future activities. Above and beyond this strength, we illustrate our model's ability to generate further insights into consumer behavioral dynamics using the video game data. We first demonstrate how individual-level parameter estimates in our model can be used to segment players based on their game-play and purchase propensities.[13] We then perform simulations and regressions to evaluate the effects of past events on players' future engagement.

6.1. Player Segmentation

We perform segmentation analysis based on each player's two intrinsic behavioral characteristics: baseline game-play and purchase propensities, which are defined as the player's expected total game-play duration and purchase count in a week[14] *without the influence of past events*. To estimate player i's baseline propensities, we set the event history $\mathcal{H}_t^i$ as empty and the representations of past influence $\{\boldsymbol{s}^i_{\text{Time}}(t), \boldsymbol{s}^i_{\text{Play}}(t), \boldsymbol{s}^i_{\text{Pur}}(t)\}$ as zero vectors. Accordingly, the ground intensity function $\lambda_i^*(t)$ in Equation (6) becomes constant with $\lambda_i^*(t) = \sigma_{\text{SP}}(b^i_{\text{Time}})$, where b^i_{Time} is the player-specific parameter in our model. Similarly, the parameters of the mark density function for player i's game-play duration and purchase count in Equations (7) to (11) are also constant over time with $\pi^i_{\text{Play}}(t) = \sigma_{\text{S}}\left(b^i_{\text{Play}}\right), \alpha^i_{\text{Play}}(t) = \sigma_{\text{SP}}\left(b^i_{\text{Play},\alpha}\right), \beta^i_{\text{Play}}(t) = \sigma_{\text{SP}}\left(b^i_{\text{Play},\beta}\right)$ and $\pi^i_{\text{Pur}}(t) = \sigma_{\text{S}}\left(b^i_{\text{Pur}}\right), r^i_{\text{Pur}}(t) = \sigma_{\text{SP}}\left(b^i_{\text{Pur},r}\right), p^i_{\text{Pur}}(t) = \sigma_{\text{SP}}\left(b^i_{\text{Pur},p}\right)$. Therefore, based on our model specification and player-specific parameter estimates, the two baseline propensities of player i can be calculated by

$$\begin{aligned} \text{Game-Play Propensity}_i &= \frac{\sigma_{\text{SP}}(b^i_{\text{Time}})\sigma_{\text{S}}(b^i_{\text{Play}})\sigma_{\text{SP}}(b^i_{\text{Play},\alpha})}{\sigma_{\text{SP}}(b^i_{\text{Play},\beta})}, \\ \text{Purchase Propensity}_i &= \frac{\sigma_{\text{SP}}(b^i_{\text{Time}})\sigma_{\text{S}}(b^i_{\text{Pur}})\sigma_{\text{SP}}(b^i_{\text{Pur},r})\sigma_{\text{S}}(b^i_{\text{Pur},p})}{\left(1-\sigma_{\text{S}}(b^i_{\text{Pur},p})\right)\left(1-\left(1-\sigma_{\text{S}}(b^i_{\text{Pur},p})\right)^{\sigma_{\text{SP}}(b^i_{\text{Pur},r})}\right)}. \end{aligned} \tag{17}$$

[13] We thank a reviewer for this valuable suggestion.

[14] We choose the time period to be a week to match the unit of time used in our model input (see Section 4).

We segment players based on the median splits of their baseline propensities[15] (0.493 hours for game-play and 1.081 counts for purchase), resulting in four segments: (1) "hardcore" players who have high propensities in both game-play and purchase; (2) "gamer" players with high game-play but low purchase propensities; (3) "buyer" players with low game-play but high purchase propensities; and (4) "casual" players who are in the lower halves for both behaviors. Table 7 provides the descriptive statistics of each segment. The average game-play propensity of the "hardcore" and "gamer" segments is almost twice that of the "buyer" and "casual" segments. Moreover, we observe the distribution of game-play propensity in the "hardcore" and "gamer" segments is more right-skewed than the distribution of purchase propensity in the "hardcore" and "buyer" segments, suggesting that players are less "extreme" in making purchases than playing the game. Interestingly, further analysis of login time reveals that the "casual" players log in more frequently on weekends than players in the other groups (0.444 vs. 0.278/0.385/0.370, all p-values < 0.05). By contrast, the "hardcore" players log in more often on weekdays. At the daily level, the "casual" players log in less frequently in the evening (from 6 p.m. to midnight) than the other three segments (0.274 vs. 0.336/0.337/0.347, all p-values < 0.01). These results generate useful insights for the video game firm seeking to understand the heterogeneity in players' game-play and purchase propensities. They also provide direct guidance on the best time frame to capture a target player segment.

6.2. Effects of Event Occurrence on Players' Future Engagement

The segmentation analyses consider players' intrinsic game-play and purchase propensities without the influence of past events. Next, we evaluate how the occurrence of past events (i.e., in-game activities and real-world sports matches) affects players' future engagement. Our procedure is

[15] We segment players using the baseline propensities estimated from the first 1, 2, 3, 4, 5, and the entire 6 months of data since the game was released, respectively. The results show that player segments are stable over time with respect to the similarity measured by the Rand index. The segmentation results reported in the paper are based on the entire 6 months of data.

Table 7 Summary Statistics of Different Player Segments

Segment	Login Time		Game-Play Propensity			Purchase Propensity		
	% of weekend logins	% of evening logins	Mean	SD	Skewness	Mean	SD	Skewness
Hardcore	0.278	0.336	0.686	0.179	1.702	1.344	0.194	1.103
Gamer	0.385	0.337	0.705	0.188	1.762	0.858	0.140	-0.299
Buyer	0.370	0.347	0.362	0.084	-0.388	1.315	0.179	1.200
Casual	0.444	0.274	0.369	0.080	-0.392	0.866	0.128	-0.219

inspired by Xu et al. (2014), who use a simulation approach to estimate the conversion effects of ad clicks. Specifically, for each in-game activity and sports match in a player's event history, we investigate the effects of its occurrence by simulating the player's subsequent behaviors in the next week under two scenarios: (1) the "actual" scenario, which includes the occurrence of the event, and (2) the "hypothetical" scenario, which assumes the event did not happen. We then examine *the effects of this event occurrence on the player's future engagement* by measuring the differences in the player's subsequent (a) login frequency, (b) game-play duration, and (c) purchase count between the two scenarios. All the details are provided in Online Appendix H.

The boxplots in Figure 4 show the distributions of the effects incurred by different event categories (i.e., in-game activity including login only, game-play only, purchase only, and concurrent game-play and purchase; and real-world sports match). Overall, despite the substantial heterogeneity in players' susceptibility to past events, the medians of their effects on players' subsequent (a) login frequency, (b) game-play duration, and (c) purchase count are all positive. Interestingly, the result suggests that the occurrence of real-world sports matches tends to have greater effects than in-game activities, and the effects incurred by login-only activities are consistently limited for future login, game-play, and purchases. Moreover, we observe that future behaviors are more affected by past behaviors of the same type; see the effects incurred by game-play only vs. purchase only activities on future (b) game-play duration and (c) purchase count. For future (c) purchase count, the occurrence of game-play only activities shows limited effects, similar to the occurrence of login-only activities.

Figure 4 **Effects of Event Occurrence on Players' Future Engagement**

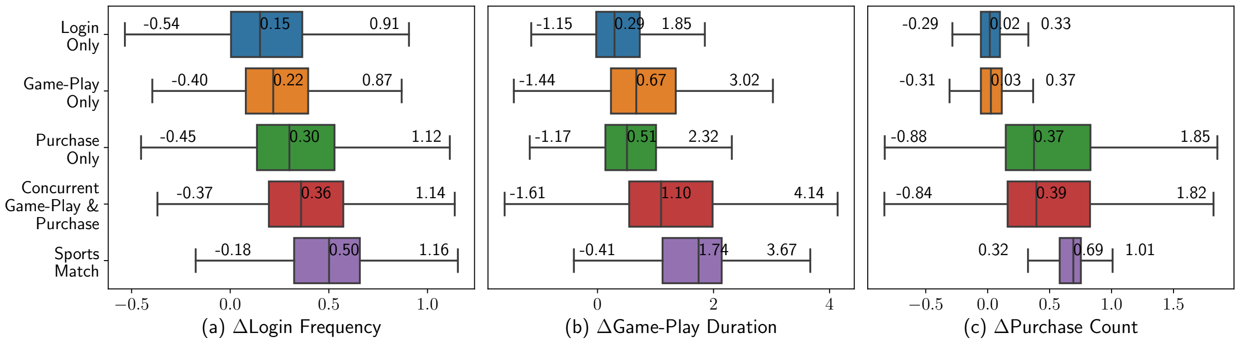
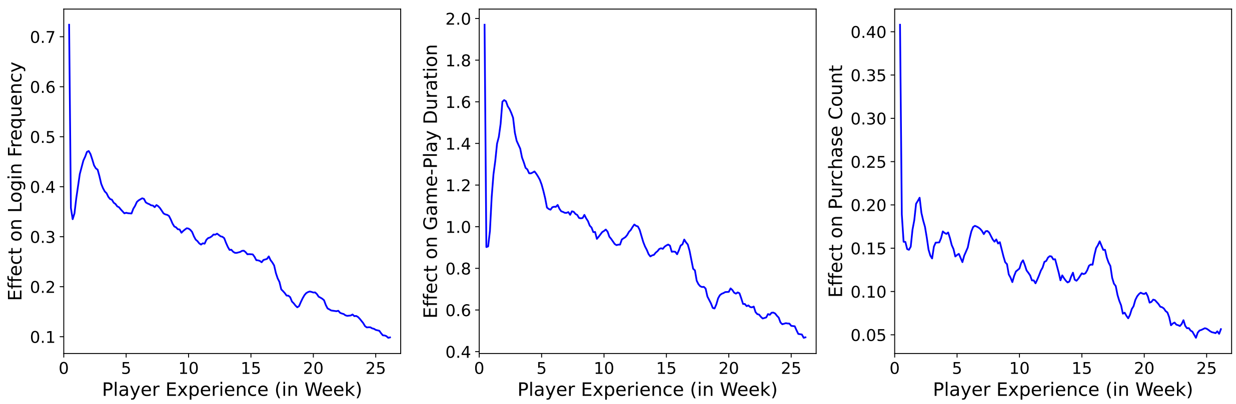
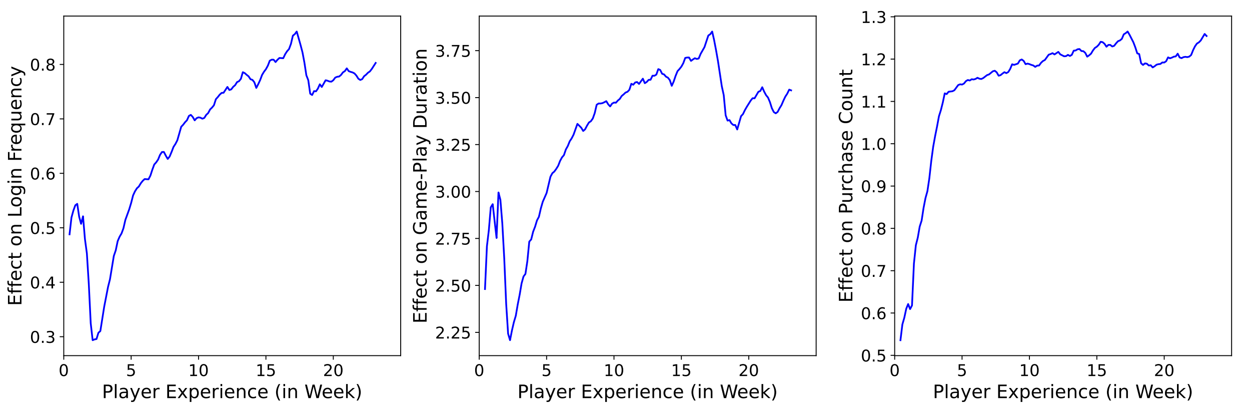


Figure 5 **Average Effects of Event Occurrence as a Function of Player Experience**



(a) Average Effects of In-Game Activities



(b) Average Effects of Real-World Sports Matches

Our simulation results also generate insights into how the effects of event occurrence vary as a function of player experience. Each player's experience is measured by the number of weeks since the player's first-time login to the game. As shown in Figure 5, the average effects of prior *in-game activities* on subsequent login frequency, game-play duration, and purchase count show a decreasing trend as players become more experienced with the video game. By contrast, the average effects incurred by *real-world sports matches* on future engagement gradually increase with the growth of player experience, indicating that more experienced players are more susceptible to the influence of real-world events. These interesting findings suggest different marketing strategies for the video game firm based on player experience. Beginners might be more curious about game features and eager to level up, so game content and design can be used to effectively motivate their future engagement. On the other hand, real-world events, such as sports matches relevant to the video game, are more appealing to experienced players and thus can be leveraged as marketing tools to engage them.

6.3. Effects of Event Marks on Players' Future Engagement

Recall that each event in a player's history is associated with a vector of marks, i.e., the game-play duration and purchase count for in-game activities, and the number of highlights and absolute score difference for sports matches. As events differ in their marks, we further conduct a set of regressions to estimate how a one-unit increase in each event mark would, on average, affect players' future engagement. The dependent variables are the outcomes generated by the simulations in Section 6.2, i.e., changes in login frequency, game-play duration, and purchase count due to an event occurrence, and the independent variables are the corresponding event marks. Fixed-effects regression models are applied to account for player heterogeneity, and we include player experience and time dummies as control variables.

Because of the page limit, we provide the details of the regression analysis in Online Appendix I and summarize the main findings here. Overall, the regression results are qualitatively consistent with our findings in Section 6.2. Players' future consumption behaviors are more affected by past behaviors of the same type, e.g., past game-play duration has a greater effect than past purchase count on future game-play, suggesting inertia in players' consumption behaviors. Moreover, in-game purchases have a larger effect than game-play on subsequent logins, but they are less motivated by past game-play, implying that more promotions are needed to foster players' in-game purchases. In terms of external stimuli, we find that the number of highlights in real-world sports matches shows stronger positive effects than the score difference on players' future engagement. Therefore, the video game firm can take this advantage to advertise in athletic fields and stadiums to attract potential players who have experienced exciting moments during the matches, or to send notifications and newsletters reminding the existing players of exciting moments from recent matches.

We also examine the effects of event marks on players' future engagement across the four player segments. Overall, the results suggest that the positive effects of past consumption behaviors are stronger among players who intrinsically like the same behavior. As such, "casual" players should be the main target group for marketing practitioners to improve their engagement, because their future game-play and purchases are relatively less susceptible to past in-game activities. The video game firm can adjust the difficulty level of the game so that "casual" players can win more easily to increase their feelings of efficacy and pleasure. The firm can also introduce more functions to raise "casual" players' curiosity to keep playing the game. To encourage them to make purchases, the firm can offer larger product discounts and promotions. In comparison, for players with high behavioral propensities, the video game firm can focus on improving their game satisfaction and customizing their in-game experiences. For example, the firm can recommend more enhancement packs to "hardcore" and "buyer" groups, or create more challenges for "hardcore" and "gamer" segments, which may also increase their need to make in-game purchases for enhancing their ability to level up. On the other hand, the effects of event marks from real-world sports matches are relatively stable across player segments. More discussions on the results and their implications can be found in Online Appendix I.

7. Discussion and Conclusions

Inspired by three unique characteristics of digital content consumption, we develop a novel attention-based neural point process approach to modeling behavioral dynamics in this fast-growing market. We illustrate the model development and applications in a major video game context as an example of digital content consumption. Our results highlight the superior capability of our model over a wide range of baselines in predicting the occurrence time, behavior combination, and consumption quantity of future player activities. The ablation studies further verify the importance of incorporating each characteristic of digital content consumption into our approach. Leveraging the individual-level parameter estimates, we further demonstrate our model's ability to segment players based on their behavioral propensities and evaluate the effects of past events on players' future engagement.

7.1. Model Contributions and Generalizability

Methodologically, our proposed model presents significant advances over existing approaches to model the complex dynamics of multi-faceted consumption behavior under external influence. Previous works either focus on modeling a single facet of consumption behaviors (e.g., Dew and Ansari 2018, Huang et al. 2019), or rely on classic multivariate point process models that make strong assumptions about interacting patterns and cannot directly model concurrent behaviors (e.g., Xu et al. 2014, Aggarwal et al. 2021). Despite the progress in the machine learning literature that makes use of more flexible RNN (e.g., Du et al. 2016, Mei and Eisner 2017), these approaches still suffer from several limitations in our application context, and our new attention-based neural point process offers substantial advantages. Instead of encoding all historical information into the last hidden state of RNN for making predictions, our novel multi-representational attention mechanism can automatically learn the relevance of all previous hidden states to form multiple representations of consumers' past events. Each representation adaptively identifies a unique and relevant aspect of the history for influencing each type of consumption behavior, thus significantly increasing the model capacity and flexibility. Moreover, our continuous-time attention mechanism can seamlessly incorporate rich elapsed irregular-time information embedded in the event history. To handle the concurrency, we use the learned representations to build a univariate point process that couples the occurrence of multiple consumption behaviors. With custom-designed hurdle models as the mark density function, our approach can efficiently handle all possible behavioral combinations and model consumption quantity of multiple behaviors without an exponential explosion in the number of model parameters. As a result, our model can be used to generate a comprehensive picture of consumer engagement in digital content consumption with individual heterogeneity.

Since our model specifications do not rely on strong assumptions about video games, the proposed approach can be readily generalized to other digital content consumption contexts or other types of marketing data with similar characteristics of consumer behaviors, namely multi-faceted and concurrent, dynamically interacting, and susceptible to external influence. For instance, our approach can be adapted to predict consumer engagement in video-streaming services, where consumers may engage in multiple behaviors concurrently (e.g., watching videos and leaving comments) and be exposed to external influences (e.g., midroll ads). Our modeling framework can also be applied to traditional marketing contexts without making *a priori* assumptions about how consumer behaviors may evolve over time. For example, our framework can be used to model the increasingly rich and multidimensional shopping behavior found in retail stores. By repeatedly observing consumers' historical visits and purchase behaviors, the store can apply our method to precisely forecast consumers' future visit time, purchase decisions, and consumption features (e.g., purchase count and shopping duration). External events, such as marketing promotions and campaigns, can be further incorporated into the model to improve the predictive performance. Under a similar framework, we can also capture other consumer decisions of interest by customizing the mark density functions, such as using a multinomial distribution to model consumers' brand choices.

7.2. Managerial Implications

In addition to the methodological contributions, our proposed approach provides an effective tool to help businesses better understand consumer behavioral dynamics in digital content consumption and has useful managerial implications.

Making efficient business plans. Our model provides a powerful tool that firms can use to forecast *when*, *what*, and *how* consumers will engage in different consumption activities. Firms can leverage this information to improve product development and marketing strategies to best capture business opportunities. For instance, based on the prediction of each player's game-play duration, firms can design and place advertisements more precisely, thereby increasing ad exposure and improving ad click-through rates. As a result, advertisers will be willing to pay a higher rate to the game platform. Knowing each player's purchase decisions can also help firms improve their content designs to generate personalized product recommendations, leading to more in-game purchases and higher revenues. Moreover, the prediction of future occurrence time offers specific advice on the timing of firms' actions. As such, firms can optimize when to push ads and make recommendations, and thus deliver their marketing designs more efficiently. In terms of operations, as consumers expect to keep a stable connection with digital content platforms, it is essential for firms to allocate the capacity to avoid service congestion or overload. Unfortunately, even leading global video-game teams, such as Blizzard's World of Warcraft, could suffer from service overload (Vaz 2020). Our model helps generate effective forecasting of consumer engagement and thus enables firms to optimize resource and capacity allocation to prevent operational problems. Having an ex-ante accurate understanding of consumers' activities, compared with ad hoc measures, is much more necessary and beneficial for improving consumer experience and satisfaction and ultimately firms' revenues.

Implementing targeted marketing. By generating individual-level parameter estimates of consumer behavioral dynamics, our model provides great opportunities for consumer segmentation and targeted marketing. As the findings of the video game data illustrate, consumers differ in game-play and purchase propensities and have varying susceptibilities to past events. Therefore, the video game firm can implement targeted marketing for different player segments. For example, the firm can adjust the difficulty level of the game and offer larger product discounts for "casual" players who are less self-motivated to stimulate their game-play and purchases. Our findings suggest that the best time window for targeting "casual" players is weekends rather than weekday evenings. Moreover, even though the average effects of past events on consumers' future engagement are positive, each consumer may have substantial heterogeneity in their susceptibility to past events. Relying on individual-level estimates of the effects, firms can detect idiosyncratic consumers whose past activities cannot motivate further engagement. Personalized interventions could then be taken to retain these consumers from fading out of digital content consumption.

Evaluating event-contingent campaigns. Our model can be used to incorporate external events and assess their effects on consumers' future engagement. In the video game example, we have demonstrated how to examine the effects of real-world sports matches on players' future game-play and in-game purchases. Note that according to our model specifications, such events need not be strictly "external"; they can be any activities outside consumers' in-consumption behaviors. For example, marketing promotions can be considered as external events affecting consumers' shopping behavior, and thus firms can apply our approach to estimate the effect of marketing promotions on each consumer's future purchase decisions; online platforms can assess the effect of midroll ads on a user's video watching, given that playing long ads may sometimes inhibit continuous watching. Based on the design of the occurrence time and event marks, e.g., the rate of a promotional discount or the length of a midroll ad, our approach provides an exciting opportunity for firms to evaluate the effectiveness of these marketing events in advance. This can provide firms with more confidence in developing and implementing marketing campaigns.

7.3. Limitations and Future Research

We conclude the paper by highlighting a few limitations and future research opportunities. First, when specifying the likelihood of our model, we assume that players in the video game act independently, given real-world events. However, consumers may communicate with each other through virtual communities, such as embedded discussion forums and third-party applications (e.g., Discord), and social interactions may potentially influence their in-consumption behaviors. While we acknowledge this data limitation, our model can be extended to account for interactions between consumers if such data becomes available. For instance, we can treat social interactions as a third event category in the observed event sequences, in addition to in-consumption activities and external events, and use the proposed RNN module to incorporate their occurrence time and marks according to Equations (1) and (2).

Second, our work provides a framework to flexibly model consumer behavioral dynamics with superior predictive power based on the data of consumers' historical behaviors and external sports matches alone. However, we acknowledge that we do not consider detailed marketing mix variables in the model because of data limitations, such as promotions and advertising. It would be interesting for future work to incorporate these marketing mix components and provide additional insights. Our modeling framework is very general and flexible to incorporate these variables, such as directly including them as event marks or allowing the individual-specific parameters $\{b^i_{\mathrm{Play}}, b^i_{\mathrm{Play},\alpha}, b^i_{\mathrm{Play},\beta}, b^i_{\mathrm{Pur}}, b^i_{\mathrm{Pur},r}, b^i_{\mathrm{Pur},p}\}$ in Equations (7) to (11) to depend on these new explanatory variables.

Finally, although our model is empirically illustrated in the video game context, the proposed modeling framework can be generalized to other digital content consumption businesses with similar characteristics, such as video-streaming services and social media platforms, or even traditional marketing contexts. Moreover, as the majority of players in our dataset have a rich event history, it would be interesting to evaluate our model in the contexts that involve relatively sparse event sequences (e.g., the purchasing of big-ticket durable goods). We leave the exploration of other contexts for future research.

Acknowledgments

The authors gratefully acknowledge the data and support from Wharton Customer Analytics (WCA). The authors thank the senior editor, the associate editor, and two anonymous reviewers for their constructive comments that have greatly improved the manuscript. The authors also thank the seminar participants at the Heinz College of Information Systems and Public Policy at Carnegie Mellon University, the China Summer Workshop on Information Management (2019), the INFORMS Data Science Workshop (2019), the Conference on Information Systems and Technology (2019), the Conference on Artificial Intelligence, Machine Learning, and Business Analytics (2019), the Workshop on Information Technologies and Systems (2021), the KDD Workshop on Machine Learning for Consumers and Markets (2021), and the Data Science in a Digital World: Methodology, Organization, and Management Workshop (2022) for helpful feedback.

Funding: Junming Yin was partly supported by Adobe Digital Experience Research Award and Amazon AWS Machine Learning Research Award (MLRA). Yue (Katherine) Feng was supported by the ECS grant from the Research Grants Council of Hong Kong [25508819]. Yong Liu gratefully acknowledges the support from Bob Eckert, Chairman of the Board at Levi Struss and former CEO and Chairman at Mattel, through the Robert A. Eckert Endowed Chair in Marketing at the University of Arizona.

References

Aalen O, Borgan O, Gjessing H (2008) *Survival and Event History Analysis: A Process Point of View* (Springer).

Abdul-Muhmin AG (2010) Repeat purchase intentions in online shopping: The role of satisfaction, attitude, and online retailers' performance. *Journal of International Consumer Marketing* 23(1):5–20.

Aggarwal V, Hwang EH, Tan Y (2021) Learning to be creative: A mutually exciting spatiotemporal point process model for idea generation in open innovation. *Information Systems Research* 32(4):1214—-1235.

Arfè A, Peluso S, Muliere P (2021) The semi-Markov beta-Stacy process: a Bayesian non-parametric prior for semi-Markov processes. *Statistical Inference for Stochastic Processes* 24(1):1–15.

Bahdanau D, Cho K, Bengio Y (2015) Neural machine translation by jointly learning to align and translate. *International Conference on Learning Representations*.

Cameron AC, Trivedi PK (2013) *Regression Analysis of Count Data* (Cambridge University Press).

Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1724–1734.

Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. *NIPS 2014 Deep Learning and Representation Learning Workshop*, URL http://arxiv.org/abs/1412.3555.

Colwell J (2007) Needs met through computer game play among adolescents. *Personality and Individual Differences* 43(8):2072–2082.

Daley DJ, Vere-Jones D (2003) *An Introduction to the Theory of Point Processes. Vol. I. Probability and Its Applications* (Springer).

Deci EL, Ryan RM (1985) *Intrinsic Motivation and Self-Determination in Human Behavior* (Springer).

Dew R, Ansari A (2018) Bayesian nonparametric customer base analysis with model-based visualizations. *Marketing Science* 37(2):216–235.

DoubleVerify (2020) Four fundamental shifts in media & advertising during 2020. https://doubleverify.com/wp-content/uploads/2020/09/DV_Four_Fundamental_Shifts_In_Media_and_Advertising_During_2020.pdf.

Du N, Dai H, Trivedi R, Upadhyay U, Gomez-Rodriguez M, Song L (2016) Recurrent marked temporal point processes: Embedding event history to vector. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1555–1564.

Erdley CA, D'Agostino PR (1988) Cognitive and affective components of automatic priming effects. *Journal of Personality and Social Psychology* 54(5):741.

Fader PS, Hardie BG, Lee KL (2005) RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research* 42(4):415–430.

Fossen BL, Schweidel DA (2019) Measuring the impact of product placement with brand-related social media conversations and website traffic. *Marketing Science* 38(3):481–499.

Goodfellow I, Bengio Y, Courville A (2016) *Deep Learning* (MIT press).

Hawkes AG (1971a) Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)* 33(3):438–443.

Hawkes AG (1971b) Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58(1):83–90.

Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Computation* 9(8):1735–1780.

Hsu MH, Chang CM, Chuang LW (2015) Understanding the determinants of online repeat purchase intention and moderating role of habit: The case of online group-buying in taiwan. *International Journal of Information Management* 35(1):45–56.

Huang Y, Jasin S, Manchanda P (2019) “Level up”: Leveraging skill and engagement to maximize player game-play in online video games. *Information Systems Research* 30(3):927–947.

Jing B (2018) Showrooming and webrooming: Information externalities between online and offline sellers. *Marketing Science* 37(3):469–483.

Jung Y, Pawlowski SD (2014) Virtual goods, real goals: Exploring means-end goal structures of consumers in social virtual worlds. *Information & Management* 51(5):520–531.

Kim Y, Ross SD (2006) An exploration of motives in sport video gaming. *International Journal of Sports Marketing and Sponsorship* 8(1):28–40.

Kingma DP, Ba J (2015) Adam: A method for stochastic optimization. *International Conference on Learning Representations*.

Lashley KS (1951) The problem of serial order in behavior. Jeffress LA, ed., *Cerebral Mechanisms in Behavior*, 112–131 (New York: Wiley).

Lin H, Sun CT (2011) Cash trade in free-to-play online games. *Games and Culture* 6(3):270–287.

Lu S, Yao D, Chen X, Grewal R (2021) Do larger audiences generate greater revenues under pay what you want? evidence from a live streaming platform. *Marketing Science* 40(5):964–984.

Mäntymäki M, Salo J (2011) Teenagers in social virtual worlds: Continuous use and purchasing behavior in habbo hotel. *Computers in Human Behavior* 27(6):2088–2097.

Martinez DM, Liu Y, Warren NL (2021) Should celebrities take sides? the impact of sociopolitical activism on spotify music streaming. Technical report.

Mei H, Eisner JM (2017) The neural Hawkes process: A neurally self-modulating multivariate point process. *Advances in Neural Information Processing Systems*, 6754–6764.

Molden DC (2014) Understanding priming effects in social psychology: An overview and integration. *Social Cognition* 32(Supplement):243–249.

Ofek E, Katona Z, Sarvary M (2011) "Bricks and Clicks": The impact of product returns on the strategies of multichannel retailers. *Marketing Science* 30(1):42–60.

Ogata Y (1981) On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory* 27(1):23–31.

Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* 77(2):257–286.

Rigat F, Muliere P (2012) Nonparametric survival regression using the beta-Stacy process. *Journal of Statistical Planning and Inference* 142(9):2688–2700.

Ryan RM, Rigby CS, Przybylski A (2006) The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion* 30(4):344–360.

Schweidel DA, Moe WW (2016) Binge watching and advertising. *Journal of Marketing* 80(5):1–19.

Seetharaman P, Chintagunta PK (2003) The proportional hazard model for purchase timing: A comparison of alternative specifications. *Journal of Business & Economic Statistics* 21(3):368–382.

Srull TK, Wyer RS (1980) Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology* 38(6):841.

Statista (2020) 2020 digital media report. https://www.statista.com/study/44526/digital-media-report/.

Tong S, Luo X, Xu B (2020) Personalized mobile marketing strategies. *Journal of the Academy of Marketing Science* 48(1):64–78.

Vaz C (2020) The broken world of warcraft servers are blizzard's "number one priority". https://www.pcgamesn.com/world-of-warcraft/shadowlands-server-problems.

Walker S, Muliere P (1997) Beta-Stacy processes and a generalization of the Pólya-urn scheme. *The Annals of Statistics* 1762–1780.

Wheeler SC, DeMarree KG, Petty RE (2014) Understanding prime-to-behavior effects: Insights from the active-self account. *Social Cognition* 32(Supplement):109–123.

Xu L, Duan JA, Whinston A (2014) Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science* 60(6):1392–1412.

Yi Y (1990) Cognitive and affective priming effects of the context for print advertisements. *Journal of advertising* 19(2):40–48.

Zhang Q, Wang W, Chen Y (2020) Frontiers: In-consumption social listening with moment-to-moment unstructured data: The case of movie appreciation and live comments. *Marketing Science* 39(2):285–295.