Counteracting Duration Bias in Video Recommendation via Counterfactual Watch Time

Haiyuan Zhao*
School of Information
Renmin University of China
Beijing, China
haiyuanzhao@ruc.edu.cn

Guohao Cai*
Huawei Noah's Ark Lab
Shenzhen, China
caiguohao1@huawei.com

Jieming Zhu
Huawei Noah's Ark Lab
Shenzhen, China
jiemingzhu@ieee.org

Zhenhua Dong
Huawei Noah's Ark Lab
Shenzhen, China
dongzhenhua@huawei.com

Jun Xu†
Gaoling School of Artificial Intelligence
Renmin University of China
Beijing, China
junxu@ruc.edu.cn

Ji-Rong Wen
Gaoling School of Artificial Intelligence
Renmin University of China
Beijing, China
jrwen@ruc.edu.cn

Abstract

In video recommendation, an ongoing effort is to satisfy users' personalized information needs by leveraging their logged watch time. However, watch time prediction suffers from duration bias, hindering its ability to reflect users' interests accurately. Existing label-correction approaches attempt to uncover user interests through grouping and normalizing observed watch time according to video duration. Although effective to some extent, we found that these approaches regard completely played records (i.e., a user watches the entire video) as equally high interest, which deviates from what we observed on real datasets: users have varied explicit feedback proportion when completely playing videos. In this paper, we introduce the *counterfactual watch time* (CWT), the potential watch time a user would spend on the video if its duration is sufficiently long. Analysis shows that the duration bias is caused by the truncation of CWT due to the video duration limitation, which usually occurs on those completely played records. Besides, a Counterfactual Watch Model (CWM) is proposed, revealing that CWT equals the time users get the maximum benefit from video recommender systems. Moreover, a cost-based transform function is defined to transform the CWT into the estimation of user interest, and the model can be learned by optimizing a counterfactual likelihood function defined over observed user watch times. Extensive experiments on three real video recommendation datasets and online A/B testing demonstrated that CWM effectively enhanced video recommendation accuracy and counteracted the duration bias.

*Both authors contributed equally to this research.

†Jun Xu is the corresponding author. Work partially done at Engineering Research Center of Next-Generation Intelligent Search and Recommendation, Ministry of Education.

CCS Concepts

• **Information systems** → **Recommender systems**.

Keywords

video recommendation, duration bias, user modelling

1 Introduction

The rising of video content platforms have attracted billions of users and become more frequent in the daily use of users nowadays [12, 15, 24, 25]. To satisfy users' information needs and enhance their engagement, developing accurate and personalized video recommender systems is critical. It is essential to incorporate various feedback signals that reflect users' interests to achieve this goal. In the video scenario, watch time has been commonly employed as an indicator of user interest and can be leveraged to enhance the accuracy of video recommender systems [5, 12].

Like other implicit feedback signals (e.g., user click) in recommender systems, directly inferring the user interest labels from the watch time is also hurt by the bias problem. One crucial bias is duration bias; that is, users' watch time is not only related to their interest but also affected by the duration (length) of the video [49, 53]. It has been observed that users naturally tend to watch for more time on longer videos, making watch time no longer a faithful reflection of the user's interest.

Existing approaches like Play Completion Rate (PCR), Watch Time Gain (WTG) [49] and Quantile-based method (D2Q) [53] regard duration bias as a problem of inconsistent watch time scales caused by the video duration. The longer the video duration, the larger the watch time scale, resulting in a longer average watch time. Therefore, these methods group and normalize the watch time according to the video duration, which keeps the watch time corresponding to different video duration consistent in scale. For
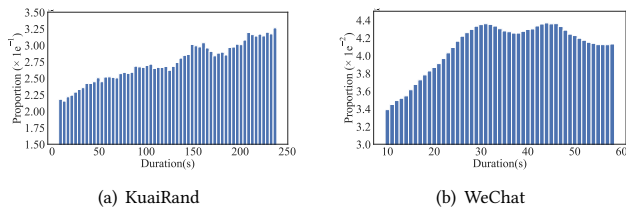
Figure 1: Users' explicit feedback proportion in completely played records grouped by video duration of (a) KuaiRand dataset (b) WeChat dataset.
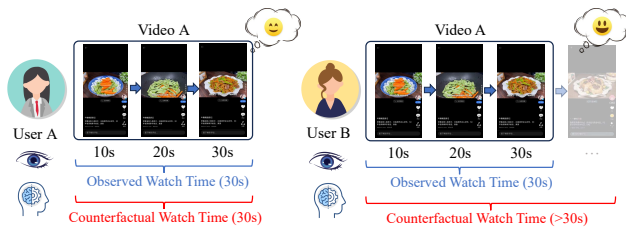

Figure 2: Comparison between counterfactual watch time and observed watch time. Users A and B have the same observed watch time but different counterfactual watch times.

example, PCR directly divides the watch time according to the corresponding video duration and then normalizes the watch time into intervals of 0 to 1. These normalized values are treated as estimated labels of user interest in downstream tasks. Since these methods mitigating bias via correcting labels, they can be categorised as **label-correction methods**. In practice, both video ranking based on user interest and accurately predicting watch time are crucial tasks. The advantage of the label-correction method is that it can not only provide an unbiased estimate of the user interest score but also predict watch time through the inverse transformation of the normalization function. In contrast, other existing methods, such as the feature-based methods: DCR [20] and CVRDD [36], primarily focus on the unbiased estimation of user interest scores. While these methods have proven effective, they struggle to simultaneously estimate both watch time and user interest scores as effectively as label-correction methods.

Although existing label-correction methods have shown effectiveness in some extent, we argue that there are still limitations. In existing methods, after the normalization, the completely played records (i.e., the user watched the whole video) are usually treated as the highest user interest, regardless of the video duration. Typically, records with explicit feedback can reflect user engagement and interest to some extent. However, from the explicit user feedback signals provided in KuaiRand dataset[1] and WeChat dataset [2], we observed that user explicit feedback in these completely played records does not align well with current methods. Specifically, Fig. 1

---

[1] https://kuairand.com/
[2] https://algo.weixin.qq.com/

shows the proportion of explicit positive feedback(i.e., like and forward) received by those completely played records, grouped by the video duration. In existing methods [49, 53], since the user has fully watched the video, these completely played records should be considered as indicating equally high interest. However, Fig. 1 clearly shows that, on both two datasets, longer completely played videos have a higher proportion of positive feedback, even though the short videos are also completely played.

The results in Fig. 1 suggest that even if users have watched the entire video, these completely played records may reflect varying levels of interest. Moreover, the shorter the video duration, the lower the user interest level that a completely played record can represent, as it corresponds to a lower explicit feedback proportion. This inspires us to derive a novel interpretation of the duration bias: *There exists a potential watch time that faithfully reflects user interest, which is truncated by the video's duration, thereby resulting in a duration bias..* As illustrated in Figure 2, let's consider two users, A and B, engaged in watching the same video with a 30s duration. Both users have completely played the video, and their watch times are recorded as 30s in the log data. However, after watching the same video, User B exhibits a higher level of interest than User A, thus appearing to be left wanting more. We then consider a counterfactual question: *How long will the user watch if the video duration is sufficiently long?* It becomes clear that User A is less likely, or even unlikely, to extend her observed watch time. The logged 30s watch time for user A can sufficiently reflect her interest. Conversely, User B prefers exceeding her observed watch time. Since her information needs are not fully satisfied after watching this video, the logged 30s watch time cannot represent B's interest.

To this end, we argue that there exists a **counterfactual watch time** (CWT) corresponding to user interest, which is partially observed due to the truncation by video duration. For incompletely played records, the CWT equals its observed watch time, thus reflecting user interest. For completely played records, the CWT could be longer than the observed watch time due to the truncation by video duration. The user's true interest level could be higher than that directly inferred from the observed watch time.

The CWT can explain the results in Fig. 1 well: various CWT are truncated by video duration in an entirely played video. Therefore, the same observed watch time may correspond to various CWT and thus cannot distinguish user interest level. Furthermore, the completion of shorter videos occurs more easily than longer ones, exacerbating the truncation of CWT. This results in a more pronounced duration bias, impacting the accuracy of interest measurement based on watch time.

To model the CWT and estimate user interests, we propose a Counterfactual Watch Model (CWM) that adopts an economic perspective to model users' CWT. Specifically, CWM treats user watching as a process of accumulating watching rewards, where the marginal rewards are indicative of user interest, and the invested watch time corresponds to the user watching cost. At the time point where a user's marginal cost equals the marginal rewards, the user attains the maximum cumulative benefit, making her actively stop watching. This time point corresponds to the aforementioned CWT. Then, a cost-based transform function is derived to transform the CWT to the estimated user interest. The duration-debiased recommendation model can be learned by optimizing a counterfactual

likelihood function defined over observed user watch times. In summary, CWM attempts to model users' consumption behavior, specifically in video scenarios. Similar to the commonly used click model [11], our CWM is beneficial for both relevance ranking and watch time prediction via effectively modeling user behaviors.

The major contributions of this work are:

(1) We provide a novel concept called *counterfactual watch time* (CWT) for interpreting the essence of duration bias;
(2) We propose a method named CWM for modeling CWT. We further develop a cost-based transform function and counterfactual likelihood function for learning a duration-debiased recommendation model;
(3) We conduct experiments on three real video recommendation datasets and online A/B testing. The result improvements demonstrate the effectiveness of our CWM.

2 Related work

Video Recommendation. In the evolution from traditional video recommendation to TV show, various methods have been developed to enhance the user experience and accuracy of recommendations. Park et al. [28] introduced a system that incorporates time factors and user preferences using 4-dimensional tensor factorization to improve recommendation accuracy. Cho et al. [10] presented a recommendation method that accounts for user feedback within specific watchable intervals to enhance user satisfaction with TV show recommendations. Qin et al. [29] proposed a model that identifies and adapts to the behavior of multiple users interacting with a TV system, thereby improving the personalization of recommendations. Most research works have transferred from traditional video to micro-video scenarios in the mobile internet era. Covington et al. [13] introduced the funnel architecture of YouTube recommender system. They predicted the expected watch time from training samples with weighted logistic regression, which utilizes observed watch time as the weight of positive samples' loss. Multi-task methods [6, 26, 35, 52] have been proposed to improve metrics such as watch time prediction, relevance of user-item pair, and number of video views together. Ma et al. [26] extend the mixture-of-experts [21] architecture to multi-gate expert knowledge integration. Tang et al. [35] proposed a shared learning structure to address the seesaw phenomenon. Chang et al. [6] proposed a plug-and-play parameter and embedding personalized network for a multi-domain and multi-task recommendation.

Counterfactual Information Retrieval. Most information retrieval systems consider users' implicit feedback as a supervision signal to infer their true interests. However, implicit feedback is influenced not only by users' interests but also by external factors. Consequently, user interest signals are often concealed within implicit feedback and remain unobserved. Researchers have drawn inspiration from causal inference techniques [46] and developed counterfactual information retrieval technology to address the biases inherent in implicit feedback. Previous work in counterfactual IR primarily focuses on mitigating position bias [1, 9, 22, 48], popularity bias [45, 50, 54] , and selection bias [32, 33, 43]. However, duration bias becomes a crucial concern when it comes to video recommendation, which has been discussed in existing studies [20, 30, 36, 49, 53]. In contrast to our approach, current methods

for correcting duration bias cannot effectively explain and eliminate duration bias.

Click model in Information Retrieval. Modeling user behaviors plays a vital role in enhancing the performance of information retrieval systems. The ability to accurately model user behaviors allows a retrieval system better to fulfill users' information needs [11]. To this end, many models have been proposed to explain or predict user click behavior in various contexts: cascade model (CM) [14], user browsing model (UBM) [16] and dynamic Bayesian network (DBN) model [7] model users' click behavior in desktop searching with different assumption; mobile click model (MCM) [27] and F-shape Click Model (FSCM) [17] further extend the understanding of users' click behaviors on mobile devices. Moreover, Borisov et al. [4] and Chen et al. [8] develop the click model into neural networks, which enable automatic dependency detection. Unlike the above click models, the UWM proposed in this paper focuses on modeling and explaining users' watching behavior since it is a better quantitative indicator of user preferences in video feeds.

3 Counterfactual watch time

In this section, we will first define the video recommendation problem and the counterfactual watch time (CWT); then we will provide supporting evidences for our proposed CWT. We also present an economic view of the user watch behavior based on the watch cost and reward. Finally, we point out the current methods' limitation from the CWT viewpoint.

3.1 Definition of counterfactual watch time

The task of user interest and watch time prediction can be formalized as follows: Suppose $\mathcal{U}$ and $\mathcal{V}$ are the sets of users and videos, respectively. We can record user $u \in \mathcal{U}$'s watching behavior on video $v \in \mathcal{V}$ as $\mathcal{D}_{u,v} = \{\mathbf{x}_{u,v}, w_{u,v}, d_v\}$, where $\mathbf{x}_{u,v} \in \mathbb{R}^k$ represents the feature vector of the sample pair and k is the feature dimension. $w_{u,v} \in \mathbb{R}^+$ denotes user u's observed watch time on video v (e.g., in seconds), while $d_v \in \mathbb{R}^+$ is the duration of video v.

Next, we will introduce a novel concept called *counterfactual watch time* CWT, which is denoted as $w_{u,v}^c$. As we have discussed before, the CWT can be defined as:

DEFINITION 1 (COUNTERFACTUAL WATCH TIME). *For user u and video v, the CWT $w_{u,v}^c \in \mathbb{R}$ is defined as the time users want to watch based on the user's interest $r_{u,v}$ if the video duration is sufficiently long. There is no correlation between $w_{u,v}^c$ and video duration d_v.*

The CWT $w_{u,v}^c$ corresponds to user interests. However, CWT $w_{u,v}^c$ does not always equal the observed watch time $w_{u,v}$ since it can be truncated by video duration d_v in practice[3]. Their relationship is formulated as follows:

$$w_{u,v} = \min(w_{u,v}^c, d_v) \iff \begin{cases} w_{u,v}^c = w_{u,v}, & \text{if } w_{u,v} < d_v; \\ w_{u,v}^c \geq w_{u,v}, & \text{if } w_{u,v} = d_v; \end{cases} \quad (1)$$

Eq. (1) indicate that observed watch time $w_{u,v}$ can be regarded as the truncated variable of CWT $w_{u,v}^c$.

---

[3]CWT may also be truncated at 0, making it difficult to discern how much a user dislikes the video. However, this study is more concerned with the videos users like than those they dislike.
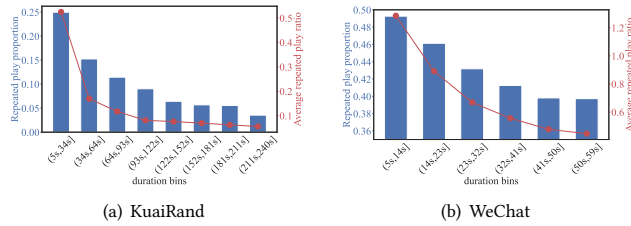
Figure 3: The repeated play proportion and average repeated play ratio in different duration bins of (a) KuaiRand dataset (b) WeChat dataset.
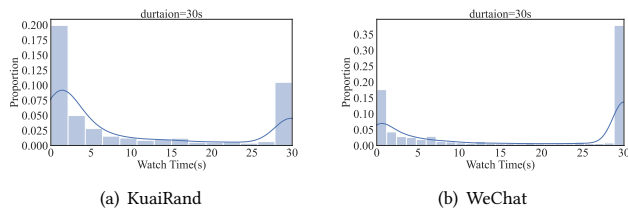

Figure 4: The bimodal distribution of watch time on (a) KuaiRand dataset (b) WeChat dataset. The video duration is 30s.
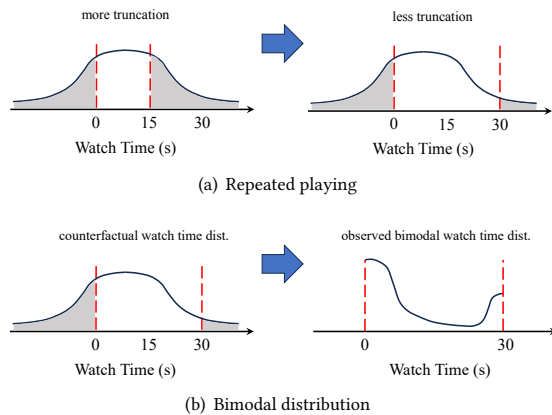

Figure 5: Employing CWT to explain (a) repeated playing (b) bimodal distribution.

3.2 The existence of counterfactual watch time

Though the CWT is not directly observable from the data, we can still find hints about the existence of CWT in real-world video recommendation datasets. Next, we will use CWT to explain two phenomena presented in real datasets: (i) users' repeated playing and (ii) the bimodal distribution of watch time.

3.2.1 Evidence 1: repeated playing. In real datasets, we found that users may engage in repeated playing (for example, by rewinding the video progress bar), leading to actual watch time that exceeds the video duration. This phenomenon is often due to users' high interest level in the current video. However, the video's duration is insufficient to meet their needs, which is similar to our definition of CWT. Since the definition of CWT necessitates a sufficiently long video duration (Definition 1), in this study, we do not equate repeat playing with CWT. In Fig. 3, we investigate the repeated playing in both the KuaiRand and WeChat datasets. Specifically, we focus on two metrics: (1) *repeated play proportion*, which represents the proportion of repeat played records within the current video duration, and (2) *average repeat play ratio*, which reflects the average extent of repeat playing within the current video duration, defined as $(w_{u,v} - d_v)/d_v$. The results depicted in Fig. 3 indicate that both the proportion of users' repeat playing and the degree of repeat playing are higher for shorter videos and decrease as video duration increases.

3.2.2 Evidence 2: bimodal distribution. Another supporting evidence is the bimodal distribution of users' watch time. For the overall distribution of all the video playing records, existing studies [47, 55] argue that logarithmic watch time obeys the Gaussian distribution. However, when we focus on a given video duration (e.g., 30s), the distribution of observed watch time turns out to be bimodal, as shown in Fig. 4. The bimodal distribution reveals that most users skip over the recommended video or completely watch it, while only a few users stop watching in the middle of the video playing [37, 51]. This abnormal distribution change is less interpreted by existing studies but can be well explained by the CWT.

3.2.3 Explanation from counterfactual watch time. Without loss of generality, we assume that CWT $w_{u,v}^c$ obeys a Gaussian distribution. However, as we have mentioned in Eq. (1), all sampled $w_{u,v}^c$ will be truncated by duration d_v. Meanwhile, $w_{u,v}^c$ will also be truncated to 0 since all recorded watch times have to be non-negative. The truncated samples are assigned to d_v or 0, respectively. As illustrated in Fig. 5(a), with the video duration increases, CWT experiences less truncation on the right side, thereby reducing the tendency for users to engage in repeat playing. Meanwhile, in Fig. 5(b), when the original Gaussian distribution is truncated, it presents a bimodal distribution, as we observed in the real-world dataset. Hence, CWT can successfully interpret the above phenomenon, which in turn supports its existence.

3.3 An economic view of user watching

To address the relationship between CWT and user interest, we model the user's watch behavior from an economic perspective. The foundation of CWT within this framework is based on the concepts of **utility maximization** [2] and **rational choice** [39]. This means users decide how much time to allocate to watching videos based on the perceived utility (satisfaction) derived from the content and their resource constraints. In accordance with these economic principles, individuals allocate resources—in this case, time—to maximize their utility. Utility can be seen as the *reward* minus the *cost*. For users, the *reward* for watching a video is the information or pleasure derived from the video content, which is determined by their interest to the video. However, users cannot earn unlimited rewards, as they face resource constraints, such as limited time. To illustrate the constraint, we introduce the concept
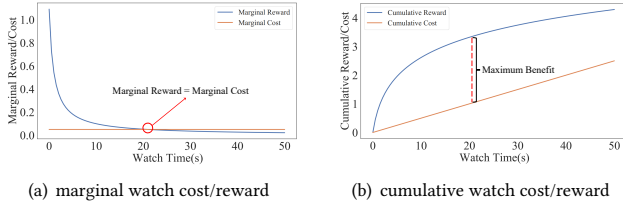
(a) marginal watch cost/reward (b) cumulative watch cost/reward

Figure 6: The economic view of user watching.

Figure 7: Computational flow of the proposed CWM.

of *watching cost*, which refers to the overall effort and resources required to watch a video, including not just the value of time for the user but also the user's mental energy and attention. This *watching cost* highlights that watching a video requires users to allocate their finite resources, which could have been spent on other activities.

In summary, We assume that the user's watching behavior follows the underlying assumptions:

- **Diminishing marginal reward**: When the user watches a video, the enjoyment or satisfaction derived per second decreases gradually, which is supported by the habituation phenomenon in the psychology filed [38] and existing practice [47]. The initial marginal watch reward corresponds to user interest.
- **Constant marginal cost**: Since the watching cost is mainly influenced by context factors, we can mildly assume that context factors are not significantly altered during the period user watching video. Therefore, we consider the marginal watch cost as a constant.
- **Rational users**: Users typically act rationally by maximizing their cumulative utility. Users will stop watching the video when the marginal watch cost equals the reward.

Based on the three assumptions above, we draw the marginal watch cost/reward curves in Fig. 6(a) and the cumulative watch cost/reward curves in Fig. 6(b). The marginal curves are obtained as derivatives of the cumulative curves. The figure shows that the marginal watch reward decreases monotonically while the marginal watch cost curve is constant. When these two curves meet, a user's marginal watch reward equals the marginal watch cost. At this point, the cumulative utility (i.e., cumulative reward minus cumulative cost) is maximized, as illustrated in Figure 6(b). Formally, we refer to this time point as our proposed CWT. Our economic view can interpret the phenomenon in Fig. 1 and Fig. 2: users interest level in a video is reflected by the time when they receives their maximum cumulative utility. When the video duration is too short to achieve each user's maximum cumulative utility point, users with either high or low interest will completely play it.

3.4 Limitation of existing methods

Finally, we point out the limitation of current methods from the viewpoint of CWT. We will prove that users' true interest cannot be inferred with only a transform function over the observed watch time, as shown in the following theorem:
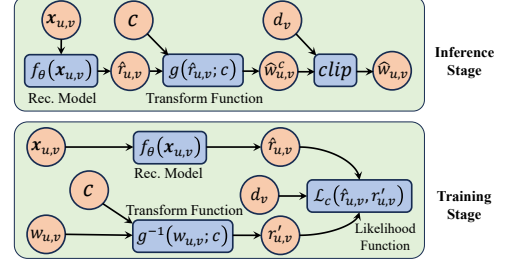
THEOREM 1 (OBSERVED WATCH TIME IS NOT THE INDICATOR OF INTEREST). *For $\forall\, \mathcal{W} \subseteq \mathbb{R}^+, g \in \mathcal{G}$, given $g : \mathcal{R} \to \mathcal{W}$, we have $\nexists\, g^{-1} : \mathcal{W} \to \mathcal{R}$, where $\mathcal{W}$ is the set of all observed watch time values, $\mathcal{G}$ is the function space, $\mathcal{R}$ is the set of all interest probability values.*

The proof of this theorem is presented in our Appendix A. This theorem indicates that existing methods fail to uncover user interest in those completely played records, especially when the video duration is short. The failure of current methods motivates us to develop a CWT-based approach to address this problem and better understand user interest.

4 Our approach

For better modeling the CWT and estimating user interest, we propose Counterfactual Watch Model (CWM). Fig 7 illustrates the flow of CWM in the inference stage and training stage. At the inference stage, a recommendation model $f_\theta(\cdot)$, parameterized with θ, estimates the user interest $\hat{r}_{u,v}$ based on the feature vector $\mathbf{x}_{u,v}$. Then a transform function $g(\cdot)$, conditioned on user watch cost c, converts the interest estimation into the CWT prediction $\hat{w}^c_{u,v}$. The actual watch time prediction $\hat{w}_{u,v}$ is obtained by truncating it through the video duration d_v.

At the training stage, to estimate the parameters θ in $f_\theta(\cdot)$, we employ a set of user activity log $\mathcal{D} \subseteq \{\mathcal{D}_{u,v} = \{\mathbf{x}_{u,v}, w_{u,v}, d_v\} : u \in \mathcal{U}, v \in \mathcal{V}\}$. For each $\mathcal{D}_{u,v} \in \mathcal{D}$, the observed watch time $w_{u,v}$ is transformed into the supervision signal of user interest by the inverse of the transform function, i.e., $r'_{u,v} = g^{-1}(w_{u,v}; c)$. Although we use observed watch time $w_{u,v}$ as the input of the inverse transform function here, we will approximate it to CWT in optimization. Then the predicted user interest $\hat{r}_{u,v} = f_\theta(\mathbf{x}_{u,v})$ is calculated with current model parameters. For all $\mathcal{D}_{u,v} \in \mathcal{D}$, comparing $r'_{u,v}$ and $\hat{r}_{u,v}$ jointly with duration d_v derives the likelihood function $\mathcal{L}_c(\cdot)$. Thus, the learning of the recommendation model can be performed by optimizing $\mathcal{L}_c(\cdot)$.

Next, we will elaborate on the design of CWM's key components: (1) the transform function $g(\cdot)$ and (2) the likelihood function $\mathcal{L}_c(\cdot)$.

4.1 Cost-based transform function

As defined in Section 3.1, the CWT $w^c_{u,v}$ reveals user interest level $r_{u,v}$. To estimate $r_{u,v}$ from $w^c_{u,v}$, we first define the transform function between $w^c_{u,v}$ and $r_{u,v}$.

Based on the economic view of the user's watching process in Section 3.3, we first formulate the cumulative watch reward and

cumulative watch cost function as:

$$F_{\text{reward}} = \omega(r_{u,v}) \cdot \log(w_{u,v}^c + 1);$$

$$F_{\text{cost}} = c \cdot w_{u,v}^c,$$

where the $\omega(r_{u,v})$ is the initial marginal reward that corresponds to user interest level $r_{u,v}$, and c is the user watch cost per second. It is evident that the derivative function of F_{reward} is monotonically decreasing while that of F_{cost} is constant, which satisfies the assumptions in Section 3.3. When their derivative values are equal, we can derive a time point (i.e., CWT) for the maximum cumulative benefit:

$$\frac{\mathrm{d}\, F_{\text{reward}}}{\mathrm{d}\, w_{u,v}^c} = \frac{\mathrm{d}\, F_{\text{cost}}}{\mathrm{d}\, w_{u,v}^c} \Rightarrow w_{u,v}^c = \frac{\omega(r_{u,v})}{c} - 1.$$

Now we will formulate the initial marginal reward function $\omega(r_{u,v})$. Since $r_{u,v} \in (0, 1)$, the $\omega(r_{u,v})$ is expected to fulfill the following conditions:(i) monotonically increasing.(ii) when $r_{u,v} \to 0$(i.e., the lowest interest level), the initial marginal reward $\omega(r_{u,v})$ should also tend to 0. (iii) when $r_{u,v} \to 1$(i.e., the highest interest level), the initial marginal reward $\omega(r_{u,v})$ should tend to positive infinity. To this end, we formulate that $\omega(r_{u,v}) = 1/(-\log r_{u,v})$, so the CWT can be further written as:

$$w_{u,v}^c = g(r_{u,v}; c) = \frac{1}{-c \log r_{u,v}} - 1. \tag{2}$$

Eq. (2) indicates how user interest $r_{u,v}$ and users' watch cost c affect their CWT. Since we aim to uncover user interest from the CWT for training recommendation model, we can rewrite Eq. (2) to its inverse function:

$$r_{u,v} = g^{-1}(w_{u,v}^c; c) = \exp\left(\frac{1}{-c(w_{u,v}^c + 1)}\right). \tag{3}$$

Both Eq. (2) and Eq. (3) are the cost-based transform functions since we introduce an extra cost parameter c (as a hyper-parameter) for controlling the conversion sensitivity from CWT to user interest or vice versa. Then we can leverage them to estimate user interest and predict the user's actual watch time.

4.2 Counterfactual likelihood function

Although we have proposed the cost-based transform function to describe the relationship between CWT and user interest, we still face the problem that CWT is truncated when the record is completely played. We need to approximate the CWT by the observed watch time to optimize our recommendation model.

4.2.1 Formulation of the counterfactual likelihood function. Inspired by the solution of survival analysis [23], we regard the observed watch time distribution as the truncated distribution of CWT. The overall likelihood function of the truncated CWT can be written as:

$$l = \prod_{w_{u,v}^c < d_v} \Pr(W^c = w_{u,v}^c, W^c < d_v \mid \mathbf{x}_{u,v}) \prod_{w_{u,v}^c \geq d_v} \Pr(W^c \geq d_v \mid \mathbf{x}_{u,v})$$

$$= \prod_{w_{u,v} < d_v} \Pr(W^c = w_{u,v}, W^c < d_v \mid \mathbf{x}_{u,v}) \prod_{w_{u,v} = d_v} \Pr(W^c \geq d_v \mid \mathbf{x}_{u,v}), \tag{4}$$

where W^c denotes the random variable of CWT. Based on Eq. (1), we can replace all $w_{u,v}^c$ by $w_{u,v}$ in the second line. Eq (4) contains two parts: when $w_{u,v} < d_v$, the likelihood function equals the joint probability that the counterfactual duration variable is equal to $w_{u,v}$

and CWT variable is not truncated. When $w_{u,v}^c \geq d_v$, the likelihood function equals the probability that the CWT variable is truncated. As discussed in Theorem 1, we cannot find a transform function for indicating user interest from observed watch time. What we know is only $w_{u,v}^c \geq d_v$, so we incorporate this prior in likelihood function, allowing the models to determine the extent to which $w_{u,v}^c$ should exceed d_v.

Maximizing Eq. (4) can reduce the duration bias caused by the truncation of CWT. To estimate the user interest, we then equally transform Eq. (4) into the likelihood function of user interest $r_{u,v}$ via the transform function in Eq. (3):

$$l = \prod_{w_{u,v} < d_v} \Pr(R = g^{-1}(w_{u,v}; c), R < g^{-1}(d_v; c) \mid \mathbf{x}_{u,v})$$

$$\times \prod_{w_{u,v} = d_v} \Pr(R \geq g^{-1}(d_v; c) \mid \mathbf{x}_{u,v})$$

$$= \prod_{w_{u,v} < d_v} \Pr(R = g^{-1}(w_{u,v}; c) \mid \mathbf{x}_{u,v}) \prod_{w_{u,v} = d_v} \Pr(R \geq g^{-1}(d_v; c) \mid \mathbf{x}_{u,v}), \tag{5}$$

where R denotes the random variable of user interest probability. Next, we will parameterize this likelihood function.

4.2.2 Parameterize and optimize the likelihood function. According to the result in [23], Eq. (5) can be parameterized with a theoretical guarantee if the random variable R obeys Gaussian distribution. For converting R into a Gaussian-distributed random variable, we employ an inverse function of the standard Gaussian-distributed cumulative distribution function:

$$g'(\cdot) = \Phi^{-1}\left[g^{-1}(\cdot)\right].$$

Then we can parameterize Eq. (5) into our counterfactual likelihood function:

$$\mathcal{L}_c = \prod_{w_{u,v} < d_v} \phi\left[\frac{g'(w_{u,v}; c) - f_\theta(\mathbf{x}_{u,v})}{\sigma}\right]$$

$$\times \prod_{w_{u,v} = d_v} \left(1 - \Phi\left[\frac{g'(d_v; c) - f_\theta(\mathbf{x}_{u,v})}{\sigma}\right]\right), \tag{6}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$[4] are the probability density function and cumulative distribution function of standard Gaussian distribution, respectively. σ is the standard deviation of the interest, which is treated as a hyper-parameter in our method. And $f_\theta(\mathbf{x}_{u,v})$ is the recommendation model for predicting user interest, its parameter is denoted as θ. Then the log-likelihood function is utilized in our training:

$$\log(\mathcal{L}_c) = \sum_{w_{u,v} < d_v} \frac{-\left(g'(w_{u,v}; c) - f_\theta(\mathbf{x}_{u,v})\right)^2}{2\sigma^2}$$

$$+ \sum_{w_{u,v} = d_v} \log \Phi\left[\frac{f_\theta(\mathbf{x}_{u,v}) - g'(d_v; c)}{\sigma}\right]. \tag{7}$$

Remark. Eq. (7) is derived from maximum likelihood estimation (MLE): The optimal parameters are those that best describe the currently observed data. For the CWT, we observe that when the video is not fully watched, the CWT equals to the actual watch time, this is the MSE part of Eq. (7). However, for the video is fully watched,

---

[4]In practice, we approximate $\Phi(\cdot)$ via Sigmoid function follow [55].

Table 1: Statistics of the datasets adopted in this study

Dataset	#Users	#Videos	#Interactions	Mean Complete Ratio(%)
KuaiRand	26,988	6,598	1,266,560	17.5%
WeChat	20,000	96,418	7,310,108	45.5%
Product	2,000,000	1,011,007	36,366,437	32.8%

the observation we only know is CWT may larger than the actual watch time, this led to the amplification part of Eq. (7). This approach to modelling truncated data is also common in endogenous problems [3] and survival analysis [40]. Although simply amplifies predictions for fully watched videos cannot precisely model interest, it can still enhance our interest predictions in a large margin, which is verified in our experimental results.

The detailed derivation from Eq. (5) to Eq. (7) can be found in our appendix. Finally, a duration-debiased recommendation model can be obtained through maximizing Eq. (7):

$$\theta^* \leftarrow \arg\max_\theta \ \log(\mathcal{L}_c).$$

4.3 Online inference

In the inference stage, given a user-video pair (u, v), the predicted interest and watch time can be calculated by the unbiased recommendation model f_{θ^*} parameterized by the learned parameter θ^*:

$$\hat{r}_{u,v} = \Phi\left(f_{\theta^*}(\mathbf{x}_{u,v})\right);$$

$$\hat{w}_{u,v} = \text{clip}\left(\frac{1}{-c\log \hat{r}_{u,v}} - 1, 0, d_v\right),$$

where $clip(x, a, b)$ function means clipping the value of x into $[a, b]$.

5 Experiments and Results

5.1 Experimental setting

We conducted experiments to verify the effectiveness of CWM on two large-scale publicly available benchmarks and a dataset collected from an industrial video product. More implementation details can be found in Appendix B. More experimental results can be found in Appendix D. The source code and dataset are available at https://github.com/hyz20/CWM.git.

5.1.1 Datasets. The experiments were conducted on two public real-world datasets: WeChat and KuaiRand. They are respectively collected from two large micro-video platforms, Wechat Channels and Kuaishou. We also conduct our evaluation in a large-scale product dataset from our video platform, which has tens of billions of daily active users. We list their statistic information in Table 1. Note that we present each dataset's completely played record percentage in the last column of Table 1. Since the duration bias usually occurs on those completely played records, their percentages in all records represent the severity of the duration bias of each dataset. According to the statistics, WeChat has the most serious bias, while KuaiRand has the least bias.

5.1.2 Evaluation. In this paper, we not only adopt our CWM for ranking videos by user interest (i.e., relevance ranking) but also for predicting users' watch time. Both tasks are of great importance in real video recommendation scenarios. As for the task of watch time prediction, we utilize users' actual watch time $w_{u,v}$ as the ground

truth, **MAE**(Mean Absolute Error) and **XAUC** [49] were used as the evaluation measures. Note that XAUC evaluates if the predictions of two samples are in the same order as their actual watch time. Such pairs are uniformly sampled, and the percentile of samples that are correctly ordered by predictions is XAUC. A larger XAUC suggests better watch time prediction performance.

As for evaluating the task of relevance ranking according to user interest, considering that the user interest labels are unobserved in real-world datasets, we defined user interests based on CWT. Given a (u, v) pair, the user interest label is defined as:

$$r_{u,v} = \begin{cases} 1 & \text{if } (d_v \leq w_{0.7} \wedge w_{u,v} = d_v) \vee (d_v > w_{0.7} \wedge w_{u,v} > w_{0.7}) \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

The $w_{0.7}$ indicates the 70% percentile of observed watch time, which is considered as the threshold for CWT. When a user watches beyond this time, we consider the user to be interested. The similar user interest definition is also adopted in [18, 51]. We will discuss the unbiasedness of $r_{u,v}$ in Appendix C. The $r_{u,v}$ is used as the ground truth for evaluating the relevance ranking task, **AUC** and **nDCG@k** are utilized as the evaluation metrics.

5.1.3 Baselines. In the experiments, we compared the proposed method with the following baselines: Three duration debiased baselines: **PCR**, **WTG** [53],**D2Q** [49], **D2Co** [51] ; Two watch time-weighted baselines: **WLR** [12] and **NDT** [47]; And a naive baseline: **VR** (value regression) which directly fit the observed watch time. For relevance ranking task, we also provide the result of **Oracle**, which is trained directly with the label defined in Eq. (8) and denote the upper bound performance of relevance ranking. Most of these methods are designed initially for relevance ranking or watch time prediction via a transform function. For the relevance ranking task, we directly rank the candidate videos via the predicted scores output by the recommendation models trained by these methods. For the watch time prediction task, we first transform the prediction of recommendation models into the interval of watch time via the inverse transform function of each method. Then we clip the estimated watch time into 0 to d_v as what we did in CWM. Two exceptions are WLR and NDT, which are implemented by a two-tower model and have no inverse transform function available. Therefore, WLR is only used in watch time prediction, and NDT is used only in relevance ranking.

To investigate the generalization of our method and the baselines, we integrate them with different backbone models. Specifically, we use recommendation models of **FM** [31], **DCN** [41] and **AutoInt** [34] as the backbone models. These three backbone models respectively represent three types of feature interactions: inner product, outer product, and attention mechanisms.

5.2 Overall performance

We compared our CWM with other baselines in the three datasets' watch time prediction task and relevance ranking task, as shown in Table 2 and Table 3, respectively. It can be seen that our CWM obtains the best performance on almost all three datasets, all backbones and both tasks significantly. We also note that on WeChat, those methods equipped with duration debiasing (e.g., WTG and D2Q) perform even worse than the naive VR method. The reason is that the WeChat dataset has much more completely played records

Table 2: The watch time prediction performance of CWM and other baselines in KuaiRand, WeChat and Product. Boldface means the best-performed methods, while <u>underline</u> means the second best-performed methods, superscripts † means the significance compared to the second best-performed methods with $p < 0.05$ of one-tailed t-test. '↓' denotes that lower is better for MAE, while higher is better for XAUC.

Dataset	Backbone Method	FM							DCN							AutoInt						
		VR	PCR	WLR	D2Q	WTG	D2Co	CWM	VR	PCR	WLR	D2Q	WTG	D2Co	CWM	VR	PCR	WLR	D2Q	WTG	D2Co	CWM
KuaiRand	MAE↓	22.100	20.974	24.279	<u>18.271</u>	23.044	22.262	**17.738**†	21.698	21.910	21.044	<u>18.131</u>	21.318	21.564	**17.420**†	21.726	21.465	23.202	<u>18.213</u>	22.771	21.826	**17.462**†
	XAUC	0.683	<u>0.697</u>	0.668	0.646	0.666	0.662	**0.714**†	0.692	<u>0.696</u>	0.680	0.649	0.695	0.693	**0.715**†	0.685	<u>0.699</u>	0.668	0.646	0.673	0.667	**0.713**†
WeChat	MAE↓	9.404	8.920	9.653	<u>8.778</u>	9.637	10.200	**8.001**†	9.317	8.709	9.149	<u>7.987</u>	8.727	8.667	**7.902**	9.405	8.910	9.815	<u>8.815</u>	9.623	10.031	**7.969**†
	XAUC	<u>0.696</u>	0.692	0.667	0.682	0.653	0.627	**0.713**†	0.703	0.717	0.684	**0.721**	0.718	0.718	0.717	<u>0.700</u>	0.693	0.661	0.679	0.653	0.634	**0.714**†
Product	MAE↓	9.411	7.395	8.913	<u>7.383</u>	7.420	7.511	**6.785**†	9.384	<u>7.320</u>	8.887	7.374	7.398	7.443	**6.648**†	9.192	<u>7.073</u>	8.668	7.169	7.125	7.293	**6.591**†
	XAUC	0.808	0.801	0.816	<u>0.817</u>	0.816	0.812	**0.833**†	0.808	0.803	0.817	<u>0.823</u>	0.820	0.817	**0.841**†	0.810	0.801	0.810	0.819	<u>0.820</u>	0.818	**0.835**†

Table 3: The relevance ranking performance of CWM and other baselines in KuaiRand, WeChat and Product.

Dataset	Backbone Method	FM								DCN								AutoInt							
		Oracle	VR	PCR	D2Q	WTG	NDT	D2Co	CWM	Oracle	VR	PCR	D2Q	WTG	NDT	D2Co	CWM	Oracle	VR	PCR	D2Q	WTG	NDT	D2Co	CWM
KuaiRand	AUC	0.738	0.661	0.686	0.679	0.684	0.668	<u>0.688</u>	**0.735**†	0.745	0.665	0.693	0.684	0.718	0.679	<u>0.730</u>	**0.735**	0.737	0.655	0.691	0.679	<u>0.692</u>	0.669	0.688	**0.734**†
	nDCG@3	0.489	0.442	<u>0.469</u>	0.464	0.462	0.461	0.464	**0.486**†	0.497	0.448	0.469	0.466	0.471	0.464	<u>0.480</u>	**0.484**	0.490	0.444	<u>0.469</u>	0.466	0.466	0.459	0.464	**0.484**†
WeChat	AUC	0.711	0.639	<u>0.651</u>	0.645	0.602	0.629	0.573	**0.703**†	0.712	0.641	0.699	0.696	<u>0.703</u>	0.672	0.702	**0.707**	0.714	0.637	0.652	0.642	0.603	0.636	0.575	**0.704**†
	nDCG@3	0.588	0.520	<u>0.540</u>	0.540	0.528	0.510	0.527	**0.581**†	0.589	0.517	<u>0.581</u>	0.577	0.576	0.527	0.570	**0.584**	0.590	0.519	<u>0.542</u>	0.540	0.534	0.512	0.532	**0.583**†
Product	AUC	0.669	0.605	0.593	0.623	0.624	<u>0.640</u>	0.629	**0.660**†	0.678	0.608	0.593	0.625	0.626	<u>0.649</u>	0.636	**0.663**†	0.672	0.610	0.593	0.629	0.631	<u>0.644</u>	0.634	**0.663**†
	nDCG@3	0.589	<u>0.556</u>	0.472	0.513	0.523	0.553	0.531	**0.582**†	0.595	<u>0.564</u>	0.475	0.516	0.533	0.563	0.540	**0.591**†	0.591	0.558	0.479	0.514	0.527	<u>0.559</u>	0.533	**0.585**†

(a) watch time prediction

(b) relevance ranking
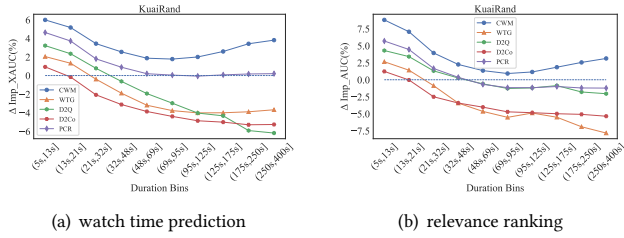
Figure 8: The relative improvement of each methods to VR on different duration bins of KuaiRand.

Figure 9: Ablation study on CWM in both KuaiRand and WeChat. Top two: watch time prediction task. Bottom two: relevance ranking task.

(45.5%) than that of KuaiRand (17.5%), showing that increasing completely played records make current debiasing methods get ineffective. In contrast, CWM improved more on WeChat than on other datasets. The results also verified the motivation of this paper: Current methods regard all completely played records as the same high interest, violating real interest distribution in real data. Therefore, when the dataset contains many completely played records (i.e., records with truncated CWT), the performance of these methods gets worse. Instead, CWM can model users' truncated CWT to estimate user interest better and predict users' actual watch time.

5.3 Effectiveness on duration debiasing

To investigate why our CWM is more effective on duration debiasing than other baselines, we divided the KuaiRand dataset into ten equal parts with different duration ranges. Then we evaluate each model on the subset of KuaiRand. The result is presented in Fig. 8. Note that the evaluation metric has different value scales among different subsets, so we report the relative improvement of each method to VR to show the extent to which these methods address duration bias. The relative improvement is $\Delta Imp = \frac{v_m - v_0}{v_0}$, where v_0 is the metric value of VR and v_m is the metric value of each method.
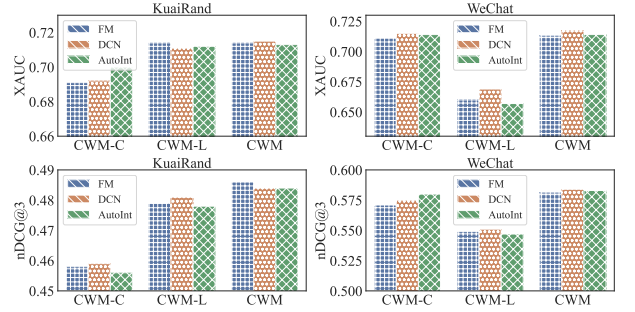
Fig 8(a) illustrates the performance of each method on the watch time prediction task, measured by ΔImp on XAUC. Fig 8(b) illustrates the performance of each method on the relevance ranking task, measured by ΔImp on AUC. In both tasks, most baselines perform better than VR in short videos (i.e., duration<30s), indicating their effectiveness on duration debiasing to some extent. However, since they simply regard completely played records as equally high interest, CWM performed better than them. We can find that these baselines perform worse in longer videos on both tasks. This is also because they regard short, completely played video recordings as high interest, leading to underestimating user interest and watch time prediction for longer videos. In contrast, CWM can model the CWT and assign fairer interest estimates to videos of different durations.

5.4 Comparison with more baselines

We have compared our CWM with other label-correction methods before, but the superiority of CWM compared with other state-of-the-art methods remains unclear. Therefore, we further implemented CVRDD [36], DCR [20], and VLDRec [30]. CVRDD treats

Table 4: The ranking performance of CWM and more state-of-the-art baselines in KuaiRand, WeChat and Product.

Dataset		KuaiRand		WeChat		Product	
Backbone	Method	AUC	nDCG@3	AUC	nDCG@3	AUC	nDCG@3
FM	Oracle	0.738	0.489	0.711	0.588	0.669	0.589
	VLDRec	0.689	0.463	0.655	0.540	0.633	0.543
	DCR	0.695	0.469	0.659	0.542	0.639	0.540
	CVRDD	0.682	0.462	0.662	0.545	0.629	0.535
	CWM	**0.735**†	**0.486**†	**0.703**†	**0.581**†	**0.660**†	**0.582**†
DCN	Oracle	0.745	0.497	0.712	0.589	0.678	0.595
	VLDRec	0.701	0.468	0.679	0.547	0.647	0.543
	DCR	0.716	0.480	0.681	0.559	0.649	0.554
	CVRDD	0.684	0.473	0.687	0.556	0.630	0.535
	CWM	**0.735**†	**0.484**	**0.707**†	**0.584**†	**0.663**†	**0.591**†
AutoInt	Oracle	0.737	0.490	0.714	0.590	0.672	0.591
	VLDRec	0.698	0.472	0.661	0.546	0.638	0.550
	DCR	0.699	0.475	0.662	0.547	0.643	0.546
	CVRDD	0.683	0.470	0.664	0.555	0.639	0.536
	CWM	**0.734**†	**0.484**	**0.704**†	**0.583**†	**0.663**†	**0.585**†

video duration as a mediation factor, DCR considers video duration as a confounder, and VLDRec employs PCR as the label with pairwise learning, all primarily feature-based or data-based approaches. In contrast, our label-correction method, CWM, consistently outperforms these baselines in terms of AUC and nDCG@3 metrics, as shown in Table 4. The superiority of CWM is consistent across all datasets and backbone models. In conclusion, our CWM method demonstrates superior performance in mitigating duration bias by modeling counterfactual watch time (CWT), highlighting its effectiveness in delivering more accurate and unbiased recommendations in short video recommendation systems.

5.5 Ablation study

We also investigate how CWM's two components benefit the CWM, i.e., the cost-based transform function and the counterfactual likelihood functions. The cost-based transform function estimates user interest from the CWT, and the counterfactual likelihood function optimizes the model unbiasedly using the observed watch time. Specifically, we produce two variants for CWM. The first is denoted as **CWM-C**, which removes the cost-based transform function and directly applies the original watch time to the counterfactual likelihood function. The second one is denoted as **CWM-L**, which replaces the counterfactual likelihood function with a mean squared error loss function.

Fig. 9 demonstrates the performance comparison between CWM and its variants on two datasets and two tasks. On KuaiRand, CWM-L obtains similar performances to CWM, while CWM-C has a significant performance drop to CWM. We argue that when CWT is less truncated (e.g., KuaiRand has only 17.5% completely played records), how it is converted into an interest estimation can primarily affect performance. On WeChat, CWM-C obtains a similar performance to CWM, while CWM-L has a significant performance drop to CWM. We argue that when CWT is heavily truncated (e.g., WeChat has 45.5% completely played records), how the observed watch time is used to approximate the learning of CWT becomes the performance bottleneck.

Table 5: Relative improvement of CWM to product baseline for one-week online A/B testing.

	MWT	VV	CTR
Product	-	-	-
CWM	+2.9%	+2.5%	+0.3%

5.6 Online A/B Testing

To verify the effectiveness of CWM in real-world recommendation scenarios, we conducted online experiments in our commercial system, a popular platform with tens of millions of active users every day. The baseline is a highly-optimized multi-task model deployed for the product. Both the baseline and CWM were trained incrementally on the same anonymous logging data, and each one serves 5% traffics, randomly selected from the same user group. As for short video recommendations, improving customers' mean watch time (MWT) is the main target. Other metrics, such as average valid viewing volume (VV) and click-through rate (CTR), are also adopted. According to the online A/B testing results shown in Table 5, we can see that CWM does help users to entertain themselves and spend more time watching the short videos.

6 Conclusion

In this study, we aim to counteract the duration bias in video recommendation. We propose counterfactual watch time (CWT) for interpreting the duration bias in video recommendation and point out that the duration bias is caused by the truncation of the user's CWT by video duration. A Counterfactual Watch Model (CWM) is then developed, revealing that the CWT equals the time users get the maximum benefit from video recommender systems. A cost-based correction function is defined to transform the CWT into the user interest, and the unbiased recommendation model can be learned by optimizing a counterfactual likelihood function defined over observed user watch times. Experimental results on three offline real datasets and online A/B testing indicate the superiority of the proposed CWM.

References

[1] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W. Bruce Croft. 2018. Unbiased Learning to Rank with Unbiased Propensity Estimation. In *The 41st International ACM SIGIR Conference on Research Development in Information Retrieval* (Ann Arbor, MI, USA) *(SIGIR '18)*. ACM, New York, NY, USA, 385–394. https://doi.org/10.1145/3209978.3209986

[2] Fuad Aleskerov, Denis Bouyssou, and Bernard Monjardet. 2007. *Utility maximization, choice and preference*. Vol. 16. Springer Science & Business Media.

[3] Takeshi Amemiya. 1984. Tobit models: A survey. *Journal of econometrics* 24, 1-2 (1984), 3–61.

[4] Alexey Borisov, Ilya Markov, Maarten de Rijke, and Pavel Serdyukov. 2016. A Neural Click Model for Web Search. In *Proceedings of the 25th International Conference on World Wide Web* (Montréal, Québec, Canada) *(WWW '16)*. International World

Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 531–541. https://doi.org/10.1145/2872427.2883033

[5] Qingpeng Cai, Zhenghai Xue, Chi Zhang, Wanqi Xue, Shuchang Liu, Ruohan Zhan, Xueliang Wang, Tianyou Zuo, Wentao Xie, Dong Zheng, Peng Jiang, and Kun Gai. 2023. Two-Stage Constrained Actor-Critic for Short Video Recommendation. In *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*. ACM, 865–875.

[6] Jianxin Chang, Chenbin Zhang, Yiqun Hui, Dewei Leng, Yanan Niu, Yang Song, and Kun Gai. 2023. PEPNet: Parameter and Embedding Personalized Network for Infusing with Personalized Prior Information. *CoRR* abs/2302.01115 (2023).

[7] Olivier Chapelle and Ya Zhang. 2009. A Dynamic Bayesian Network Click Model for Web Search Ranking. In *Proceedings of the 18th International Conference on World Wide Web* (Madrid, Spain) *(WWW '09)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/1526709.1526711

[8] Jia Chen, Jiaxin Mao, Yiqun Liu, Min Zhang, and Shaoping Ma. 2020. A Context-Aware Click Model for Web Search. In *Proceedings of the 13th International Conference on Web Search and Data Mining* (Houston, TX, USA) *(WSDM '20)*. Association for Computing Machinery, New York, NY, USA, 88–96. https://doi.org/10.1145/3336191.3371819

[9] Mouxiang Chen, Chenghao Liu, Jianling Sun, and Steven C.H. Hoi. 2021. Adapting Interactional Observation Embedding for Counterfactual Learning to Rank. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, Canada) *(SIGIR '21)*. Association for Computing Machinery, New York, NY, USA, 285–294. https://doi.org/10.1145/3404835.3462901

[10] Kyung-Jae Cho, Yeon-Chang Lee, Kyungsik Han, Jaeho Choi, and Sang-Wook Kim. 2019. No, that's not my feedback: TV show recommendation using watchable interval. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. IEEE, 316–327.

[11] Aleksandr Chuklin, Ilya Markov, and Maarten de Rijke. 2016. Click Models for Web Search and Their Applications to IR: WSDM 2016 Tutorial. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining* (San Francisco, California, USA) *(WSDM '16)*. Association for Computing Machinery, New York, NY, USA, 689–690. https://doi.org/10.1145/2835776.2855113

[12] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems* (Boston, Massachusetts, USA) *(RecSys '16)*. Association for Computing Machinery, New York, NY, USA, 191–198. https://doi.org/10.1145/2959100.2959190

[13] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, September 15-19, 2016*. ACM, 191–198.

[14] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. 2008. An Experimental Comparison of Click Position-Bias Models. In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (Palo Alto, California, USA) *(WSDM '08)*. Association for Computing Machinery, New York, NY, USA, 87–94. https://doi.org/10.1145/1341531.1341545

[15] James Davidson, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi, Sujoy Gupta, Yu He, Mike Lambert, Blake Livingston, and Dasarathi Sampath. 2010. The YouTube Video Recommendation System. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (Barcelona, Spain) *(RecSys '10)*. Association for Computing Machinery, New York, NY, USA, 293–296. https://doi.org/10.1145/1864708.1864770

[16] Georges E. Dupret and Benjamin Piwowarski. 2008. A User Browsing Model to Predict Search Engine Click Data from Past Observations.. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (Singapore, Singapore) *(SIGIR '08)*. Association for Computing Machinery, New York, NY, USA, 331–338. https://doi.org/10.1145/1390334.1390392

[17] Lingyue Fu, Jianghao Lin, Weiwen Liu, Ruiming Tang, Weinan Zhang, Rui Zhang, and Yong Yu. 2023. An F-Shape Click Model for Information Retrieval on Multi-Block Mobile Pages. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining* (Singapore, Singapore) *(WSDM '23)*. Association for Computing Machinery, New York, NY, USA, 1057–1065. https://doi.org/10.1145/3539597.3570365

[18] Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. 2022. KuaiRand: An Unbiased Sequential Recommendation Dataset with Randomly Exposed Videos. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management* (Atlanta, GA, USA) *(CIKM '22)*. 5 pages. https://doi.org/10.1145/3511808.3557624

[19] Yunjun Gao, Yuntao Du, Yujia Hu, Lu Chen, Xinjun Zhu, Ziquan Fang, and Baihua Zheng. 2022. Self-Guided Learning to Denoise for Robust Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Madrid, Spain) *(SIGIR '22)*. Association for Computing Machinery, New York, NY, USA, 1412–1422. https://doi.org/10.1145/3477495.3532059

[20] Xiangnan He, Yang Zhang, Fuli Feng, Chonggang Song, Lingling Yi, Guohui Ling, and Yongdong Zhang. 2023. Addressing Confounding Feature Issue for Causal Recommendation. *ACM Trans. Inf. Syst.* 41, 3, Article 53 (feb 2023), 23 pages. https://doi.org/10.1145/3559757

[21] Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. 1991. Adaptive Mixtures of Local Experts. *Neural Comput.* 3, 1 (1991), 79–87.

[22] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. 2017. Unbiased Learning-to-Rank with Biased Feedback. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (Cambridge, United Kingdom) *(WSDM '17)*. ACM, New York, NY, USA, 781–789. https://doi.org/10.1145/3018661.3018699

[23] Yan Li, Kevin S. Xu, and Chandan K. Reddy. 2016. Regularized Parametric Regression for High-dimensional Survival Analysis. In *Proceedings of the 2016 SIAM International Conference on Data Mining, Miami, Florida, USA, May 5-7, 2016*, Sanjay Chawla Venkatasubramanian and Wagner Meira Jr. (Eds.). SIAM, 765–773. https://doi.org/10.1137/1.9781611974348.86

[24] Shang Liu, Zhenzhong Chen, Hongyi Liu, and Xinghai Hu. 2019. User-Video Co-Attention Network for Personalized Micro-Video Recommendation. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 3020–3026. https://doi.org/10.1145/3308558.3313513

[25] Yiyu Liu, Qian Liu, Yu Tian, Changping Wang, Yanan Niu, Yang Song, and Chenliang Li. 2021. Concept-Aware Denoising Graph Neural Network for Micro-Video Recommendation. In *Proceedings of the 30th ACM International Conference on Information Knowledge Management* (Virtual Event, Queensland, Australia) *(CIKM '21)*. Association for Computing Machinery, New York, NY, USA, 1099–1108. https://doi.org/10.1145/3459637.3482417

[26] Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H. Chi. 2018. Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*. ACM, 1930–1939.

[27] Jiaxin Mao, Cheng Luo, Min Zhang, and Shaoping Ma. 2018. Constructing Click Models for Mobile Search. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (Ann Arbor, MI, USA) *(SIGIR '18)*. Association for Computing Machinery, New York, NY, USA, 775–784. https://doi.org/10.1145/3209978.3210060

[28] Yoojin Park, Jinoh Oh, and Hwanjo Yu. 2017. RecTime: Real-time recommender system for online broadcasting. *Information Sciences* 409 (2017), 1–16.

[29] Jiarui Qin, Jiachen Zhu, Yankai Liu, Junchao Gao, Jianjie Ying, Chaoxiong Liu, Ding Wang, Junlan Feng, Chao Deng, Xiaozheng Wang, et al. 2023. Learning to distinguish multi-user coupling behaviors for tv recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*. 204–212.

[30] Yuhan Quan, Jingtao Ding, Chen Gao, Nian Li, Lingling Yi, Depeng Jin, and Yong Li. 2023. Alleviating Video-length Effect for Micro-video Recommendation. *ACM Trans. Inf. Syst.* 42, 2, Article 44 (nov 2023), 24 pages. https://doi.org/10.1145/3617826

[31] Steffen Rendle. 2012. Factorization Machines with LibFM. *ACM Trans. Intell. Syst. Technol.* 3, 3, Article 57 (may 2012), 22 pages. https://doi.org/10.1145/2168752.2168771

[32] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased Recommender Learning from Missing-Not-At-Random Implicit Feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining* (Houston, TX, USA) *(WSDM '20)*. Association for Computing Machinery, New York, NY, USA, 501–509. https://doi.org/10.1145/3336191.3371783

[33] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*. PMLR, 1670–1679.

[34] Weiping Song, Chence Shi, Zhiping Xiao, Zhijian Duan, Yewen Xu, Ming Zhang, and Jian Tang. 2019. AutoInt: Automatic Feature Interaction Learning via Self-Attentive Neural Networks. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (Beijing, China) *(CIKM '19)*. Association for Computing Machinery, New York, NY, USA, 1161–1170. https://doi.org/10.1145/3357384.3357925

[35] Hongyan Tang, Junning Liu, Ming Zhao, and Xudong Gong. 2020. Progressive Layered Extraction (PLE): A Novel Multi-Task Learning (MTL) Model for Personalized Recommendations. In *RecSys 2020: Fourteenth ACM Conference on Recommender Systems, Virtual Event, Brazil, September 22-26, 2020*. ACM, 269–278.

[36] Shisong Tang, Qing Li, Dingmin Wang, Ci Gao, Wentao Xiao, Dan Zhao, Yong Jiang, Qian Ma, and Aoyang Zhang. 2023. Counterfactual Video Recommendation for Duration Debiasing. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (<conf-loc>, <city>Long Beach</city>, <state>CA</state>, <country>USA</country>, </conf-loc>) *(KDD '23)*. Association for Computing Machinery, New York, NY, USA, 4894–4903. https://doi.org/10.1145/3580305.3599797

[37] Shisong Tang, Qing Li, Dingmin Wang, Ci Gao, Wentao Xiao, Dan Zhao, Yong Jiang, Qian Ma, and Aoyang Zhang. 2023. Counterfactual Video Recommendation for Duration Debiasing. In *Proceedings of the 29th ACM SIGKDD Conference*

on Knowledge Discovery and Data Mining (Long Beach, CA, USA) *(KDD '23)*. Association for Computing Machinery, New York, NY, USA, 4894–4903. https://doi.org/10.1145/3580305.3599797

[38] Richard F Thompson and William A Spencer. 1966. Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychological review* 73, 1 (1966), 16.

[39] Nicolaas J Vriend. 1996. Rational behavior and economic theory. *Journal of Economic Behavior & Organization* 29, 2 (1996), 263–285.

[40] Ping Wang, Yan Li, and Chandan K Reddy. 2019. Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)* 51, 6 (2019), 1–36.

[41] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. 2017. Deep & Cross Network for Ad Click Predictions. In *Proceedings of the ADKDD'17* (Halifax, NS, Canada) *(ADKDD'17)*. Association for Computing Machinery, New York, NY, USA, Article 12, 7 pages. https://doi.org/10.1145/3124749.3124754

[42] Wenjie Wang, Fuli Feng, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. 2021. Denoising Implicit Feedback for Recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (Virtual Event, Israel) *(WSDM '21)*. Association for Computing Machinery, New York, NY, USA, 373–381. https://doi.org/10.1145/3437963.3441800

[43] Xuanhui Wang, Michael Bendersky, Donald Metzler, and Marc Najork. 2016. Learning to Rank with Selection Bias in Personal Search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Pisa, Italy) *(SIGIR '16)*. ACM, New York, NY, USA, 115–124.

[44] Yu Wang, Xin Xin, Zaiqiao Meng, Joemon M Jose, Fuli Feng, and Xiangnan He. 2022. Learning Robust Recommenders through Cross-Model Agreement. In *Proceedings of the ACM Web Conference 2022* (Virtual Event, Lyon, France) *(WWW '22)*. Association for Computing Machinery, New York, NY, USA, 2015–2025. https://doi.org/10.1145/3485447.3512202

[45] Tianxin Wei, Fuli Feng, Jiawei Chen, Ziwei Wu, Jinfeng Yi, and Xiangnan He. 2021. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery Data Mining* (Virtual Event, Singapore) *(KDD '21)*. Association for Computing Machinery, New York, NY, USA, 1791–1800. https://doi.org/10.1145/3447548.3467289

[46] Peng Wu, Haoxuan Li, Yuhao Deng, Wenjie Hu, Quanyu Dai, Zhenhua Dong, Jie Sun, Rui Zhang, and Xiao-Hua Zhou. 2022. On the Opportunity of Causal Learning in Recommendation Systems: Foundation, Estimation, Prediction and Challenges. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*. 5646–5653. Survey Track.

[47] Ruobing Xie, Lin Ma, Shaoliang Zhang, Feng Xia, and Leyu Lin. 2023. Reweighting Clicks with Dwell Time in Recommendation. In *Companion Proceedings of the ACM Web Conference 2023* (Austin, TX, USA) *(WWW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 341–345. https://doi.org/10.1145/3543873.3584624

[48] Bowen Yuan, Yaxu Liu, Jui-Yang Hsia, Zhenhua Dong, and Chih-Jen Lin. 2020. Unbiased Ad Click Prediction for Position-Aware Advertising Systems. In *Fourteenth ACM Conference on Recommender Systems* (Virtual Event, Brazil) *(RecSys '20)*. ACM, New York, NY, USA, 368–377. https://doi.org/10.1145/3383313.3412241

[49] Ruohan Zhan, Changhua Pei, Qiang Su, Jianfeng Wen, Xueliang Wang, Guanyu Mu, Dong Zheng, Peng Jiang, and Kun Gai. 2022. Deconfounding Duration Bias in Watch-Time Prediction for Video Recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (Washington DC, USA) *(KDD '22)*. Association for Computing Machinery, New York, NY, USA, 4472–4481. https://doi.org/10.1145/3534678.3539092

[50] Yang Zhang, Fuli Feng, Xiangnan He, Tianxin Wei, Chonggang Song, Guohui Ling, and Yongdong Zhang. 2021. Causal Intervention for Leveraging Popularity Bias in Recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, Canada) *(SIGIR '21)*. ACM, New York, NY, USA, 11–20. https://doi.org/10.1145/3404835.3462875

[51] Haiyuan Zhao, Lei Zhang, Jun Xu, Guohao Cai, Zhenhua Dong, and Ji-Rong Wen. 2023. Uncovering User Interest from Biased and Noised Watch Time in Video Recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems* (Singapore, Singapore) *(RecSys '23)*. Association for Computing Machinery, New York, NY, USA, 528–539. https://doi.org/10.1145/3604915.3608797

[52] Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed H. Chi. 2019. Recommending what video to watch next: a multitask ranking system. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019*. ACM, 43–51.

[53] Yu Zheng, Chen Gao, Jingtao Ding, Lingling Yi, Depeng Jin, Yong Li, and Meng Wang. 2022. DVR: Micro-Video Recommendation Optimizing Watch-Time-Gain under Duration Bias. In *Proceedings of the 30th ACM International Conference on Multimedia* (Lisboa, Portugal) *(MM '22)*. Association for Computing Machinery, New York, NY, USA, 334–345. https://doi.org/10.1145/3503161.3548428

[54] Yu Zheng, Chen Gao, Xiang Li, Xiangnan He, Yong Li, and Depeng Jin. 2021. Disentangling User Interest and Conformity for Recommendation with Causal Embedding. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*.

ACM, New York, NY, USA, 2980–2991. https://doi.org/10.1145/3442381.3449788

[55] Tengfei Zhou, Hui Qian, Zebang Shen, Chao Zhang, Chengwei Wang, Shichen Liu, and Wenwu Ou. 2018. JUMP: a joint predictor for user click and dwell time. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. AAAI Press. 3704–3710.

A Proof of theorem 1

PROOF. As we have analyzed before, the CWT $w_{u,v}^c$ corresponds to user interests $r_{u,v}$, i.e., $w_{u,v}^c = g(r_{u,v})$. Based on Eq. (1), we can obtain the functions of $w_{u,v}$ on $r_{u,v}$:

$$w_{u,v} = \min(g(r_{u,v}), d_v), \quad r_{u,v} \in \mathcal{R}, \ w_{u,v} \in \mathcal{W}$$

Next, we need to prove that the above function does not always have an inverse function, w.r.t any $w_{u,v}$. Similar to Eq. (1), we rewrite the above function as a segmented function:

$$r_{u,v} = g^{-1}(w_{u,v}), \quad \text{if } w_{u,v} < d_v;$$
$$r_{u,v} \geq g^{-1}(w_{u,v}), \quad \text{if } w_{u,v} = d_v;$$

Note that when $w_{u,v} = d_v$ (i.e., completely play a video), we can only obtain an inequality between $w_{u,v}$ and $r_{u,v}$, thus proving that there is no such an inverse function $r_{u,v} = g^{-1}(w_{u,v})$ for all $w_{u,v} \in \mathcal{W}$. □

B Detailed Experimental Setting

In our paper, we briefly described the dataset used in our papers and the baselines due to the limitation of pages. Now, we put a more detailed experimental setting in this appendix.

WeChat. This dataset is released by WeChat Big Data Challenge 2021, containing the WeChat Channels logs within two weeks. Following the practice in [53], we split the data into the first 10 days, the middle 2 days, and the last 2 days as training, validation, and test set. We noticed that there is an unusually high proportion of 60s videos among the dataset (17.3%), so our experiments were conducted on a subset of 60s videos that were excluded (i.e., the duration range is [5s,59s] in the subset). The adopted input features include *userid,feedid,device,authorid,bgm_song_id,bgm_singer_id,user_type*.

KuaiRand [18]. KuaiRand is a newly released sequential recommendation dataset collected from KuaiShou. As suggested in [18], we utilized one of the subsets *KuaiRand-pure* in this study. To mitigate the sparsity problem, we selected data from which the video duration is up to 400s. We split the data into the first 14 days, the middle 7 days, and the last 10 days as training, validation, and test set. The adopted input features include *user_id, video_id, author_id, music_id, follow_user_num_range,register_days_range, fans_user_num_range, friend_user_num_range, user_active_degree, most_popular_tag, video_type,upload_type,tab*.

Product. We collect the product dataset from the server log of our video platform, which samples the log data from June 19, 2023 to June 25, 2023. Due to the data imbalance, we intercepted the records below 42 seconds as the final training set. We trained our model on this dataset and tested them in the last two hours of server log data. The pretrained ID embedding and side information are used as feature inputs for all methods.

In this study, we implement WLR following the details in [49]. For D2Q, the group number is set to 60 in KuaiRand, 30 in WeChat, and 10 in our Product dataset. For NDT, we set its hyper-parameters as the author suggested in their paper [47]. For D2Co, the window size and sensitivity-controlled term are set as suggested in their
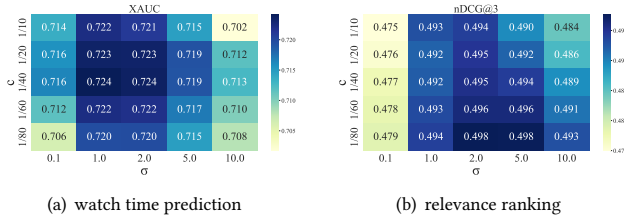
(a) watch time prediction

(b) relevance ranking

Figure 11: The parameter sensitivity of CWM on KuaiRand dataset where the backbone model is FM.

(a) Duration≤100s

(b) Duration=30s
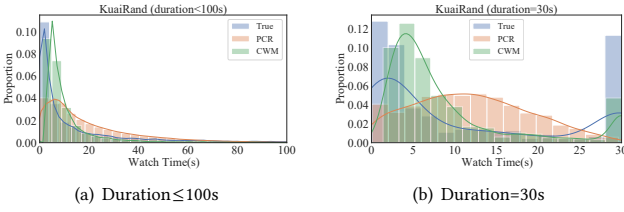
Figure 12: The true and estimated watch time distribution on KuaiRand dataset.
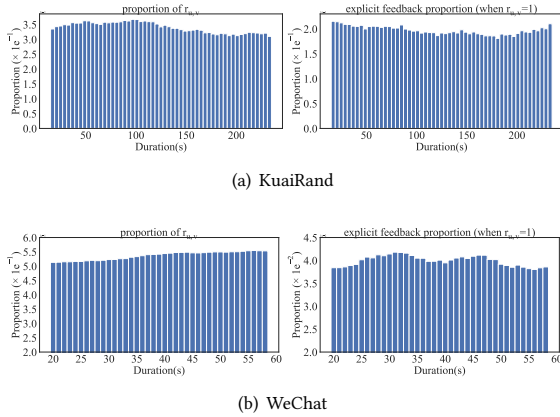

(a) KuaiRand

(b) WeChat

Figure 10: (Left) The proportion of user interest labels $r_{u,v}$ and (Right) The proportion of explicit feedback when $r_{u,v} = 1$ grouped by video duration on (a) KuaiRand and (b) WeChat.

paper [51]. We utilize Adam as the optimizer and set the initial learning rate as $5e^{-4}$ for all methods. The batch size is set as 512. For all the backbone models, we set their latent embedding dimension to 10. For all methods with neural networks, the hidden units are set to 64 while the dropout ratio is set to 0.2. The value of user cost c and σ in our CWM is set to $(1/40, 2)$ in the KuaiRand dataset, $(1/40, 20)$ in the WeChat dataset, and $(1/5, 5)$ in the Product dataset. We tune our hyperparameters on the validation set while evaluating the performance on the test set.

C The Unbiasedness of Interest Labels

To evaluate the performance of relevance ranking tasks in this study, we need an unbiased indicator of user interest first. However, the

user interest labels are unobserved in real-world datasets. Although explicit feedback can reflect user interest and is not affected by duration bias, it suffers from severe selection bias and noise [19, 42, 44]: this indicates that users might not provide explicit feedback for videos they like and might mistakenly provide explicit feedback for videos they dislike. Therefore, using explicit feedback as a label for evaluating relevance ranking tasks is inappropriate. To this end, we defined user interests based on CWT in Eq. (8). However, the unbiasedness of this interest label still needs to be discussed.

To achieve duration unbiased, the user interest indicator $r_{u,v}$ needs to fulfill two characteristics: (1) it should be independent of video duration, and (2) when $r_{u,v} = 1$, user explicit feedback should be equivalent across all video duration (thus mitigating the issue described in Fig 1). To verify whether our interest indicator defined in Eq. (8) satisfies the above characteristics, we calculate the proportion of user interest labels $r_{u,v}$ and the proportion of explicit feedback when $r_{u,v} = 1$ grouped by video duration. Note that since videos shorter than $w_{0.7}$ cannot achieve a watch time of $w_{0.7}$, we exclude these videos from our analysis. The results are presented in Fig 10. It is evident that our defined $r_{u,v}$ roughly satisfies the above two characteristics on both the KuaiRand and WeChat datasets, thus indicating its unbiasedness.

D More Experimental results

D.1 Parameter sensitivity

There are two hyper-parameters in the proposed CWM: one is the user watch cost c in the cost-based transform function (Eq. (3)). The larger the c, the more sensitive users are to watch time; Another is the variance term σ of user interest in counterfactual likelihood function (Eq (6)). The larger the value of σ, the more dispersed the user's interest distribution is. Fig. 11 illustrates the performance changes of recommendation with different values of c and σ. For watch time prediction (Fig. 11(a)), the best hyper-parameter is $\sigma \in (1.0, 2.0) \wedge c \in (1/40, 1/20)$; For relevance ranking (Fig. 11(b)), the best hyper-parameters is $\sigma \in (2.0, 5.0) \wedge c \in (1/80, 1/60)$. Note that the best hyper-parameters of two tasks may not be the same. In practice, it is necessary to adjust the hyper-parameters to make CWM perform best.

D.2 Better fit to the true watch time distribution

We examine whether CWM can fit the true watch time distribution. As a comparison, we also present the estimated watch time distribution by PCR, which is representative of existing debiasing methods. Fig 12(a) shows the true and estimated watch time distribution on videos with less than 100s duration on KuaiRand. We can find that the estimated watch time distribution by PCR is more flattening than the true distribution. It overestimates higher watch time (i.e., >10s). In contrast, our CWM can fit the true distribution even better. Fig 12(b) shows the true and estimated watch time distribution on videos with 30s duration on KuaiRand. It can be found that PCR only estimates a single-peaked distribution which differs significantly from the true bimodal distribution. In comparison, CWM can estimate a similar bimodal distribution to the true distribution, demonstrating CWM's effectiveness in watch time prediction.