TCFIMT: TEMPORAL COUNTERFACTUAL FORECASTING FROM INDIVIDUAL MULTIPLE TREATMENT PERSPECTIVE

Anonymous authors
Paper under double-blind review

ABSTRACT

Determining causal effects of temporal multi-intervention assists decision-making. Restricted by time-varying bias, selection bias, and interactions of multiple interventions, the disentanglement and estimation of multiple treatment effects from individual temporal data is still rare. To tackle these challenges, we propose a comprehensive framework of temporal counterfactual forecasting based on balanced representation from an individual multiple treatment perspective (TCFimt). TCFimt constructs adversarial tasks in a seq2seq framework to alleviate selection and time-varying bias and designs a contrastive learning-based block to decouple a mixed treatment effect into separated main treatment effects and causal interactions which further improves estimation accuracy. Through implementing experiments on two real-world datasets from distinct fields, the proposed method shows satisfactory performance in predicting future outcomes with specific treatments and in choosing optimal treatment type and timing than state-of-the-art methods.

1 INTRODUCTION

Causal analysis in temporal scenarios is to explore which factors lead to the results in the future, which could assist in decision-making. In real-life scenarios, decision-makers are often faced with the dilemma of which decision or even a combination of options to choose, hence it is particularly important for them to reliably estimate the effects of distinct interventions and intervention interactions. For example, in the game field, game companies are really concerned about the changing revenue, which reflects various states of a game. As shown on the left of Figure 1, when a noticeable change in a game indicator is observed, which is due to several mixed mega-events (e.g., game anniversary, the newly opened game server, and new gameplay) going on during this period. To weigh costs and benefits, they are eager to know what will happen if adopting other mixed interventions. Similarly, in the field of health care, doctors pay more attention to changes in vital signs, leading to what treatment regimens are needed. However, due to the physical characteristics of different patients, the treatment plans should be varied. The doctors also want to know what will happen in if the treatment is changed (right panel of Figure 1). The above questions motivate our research on counterfactual forecasting of mixed interventions on time-series data in this paper.

Existing causal inference techniques under the longitudinal setting such as Recurrent Marginal Structural Networks (RMSN, Lim et al., 2018) and Counterfactual Recurrent Network (CRN, Bica et al., 2020) cannot be applied to conduct causal inference of multiple treatments in temporal data. First of all, only one of the multiple treatment result (factual outcome) can be observed, while the results of other interventions at that moment (counterfactual outcomes) are not available. This makes us never obtain the entire vector of outcomes. Second, the observational data is prone to multiple treatment selection biases. For example, it is easier for the rich to get better treatment, which causes the distribution of features among patients to vary drastically across different choices of treatments. Third, these observed outputs are the result of mixed effects of multiple interventions, making it difficult to estimate the effects of a single intervention and the causal interactions.

Accordingly, to solve the above-mentioned challenges, this paper designs a Temporal Counterfactual Forecasting network from an Individual Multiple Treatment perspective (TCFimt), for enhancing prediction accuracy and inference effectiveness. Different from causal graphs or do-calculus, TC-

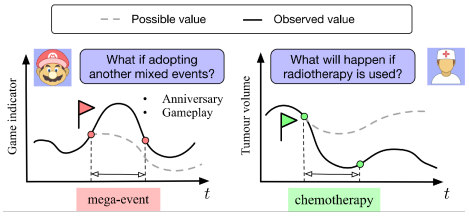


Figure 1: Counterfactual forecasting in real-world cases.

Fimt can effectively estimate multiple treatment effects for individual levels without graph search. in Specifically, we consider the following settings to address these issues: 1) For unobserved counterfactual data, we design a corruption function to generate the pseudo-counterfactual data of multiple treatments. 2) Moreover, by simultaneously maximizing the error of the treatment classification task and minimizing the error of the outcomes prediction task, the selection bias and time-varying bias are alleviated by employing an adversarial training way. 3) Finally, an effect disentanglement block based on contrastive learning is dedicated to further enhancing the prediction of effects under mixed interventions. The contributions of this research are summarized as follows:

- To our best knowledge, our paper is the first to explore the interventional effect decoupling and causal interaction issue in temporal data, which is an attempt at causal inference in complex scenarios.
- The proposed TCFimt method combines the way of adversarial training and contrastive learning, which is beneficial to solving the issues of temporal counterfactual forecasting with mixed interventions in one treatment plan.
- We provide a theoretical analysis of algorithms for learning the balanced representation in temporal individual multiple treatment effects estimation.
- The implemented experiments on real-world data in two different fields validate the effectiveness of our methods.

2 PRELIMINARY

In this section, we will employ the following notations[1]. For time any $t > 0$, we denote time-varying feature variables by $\mathbf{X}_t \in \mathcal{X} \subset \mathbb{R}^{D_x}$, a binary time-dependent treatment action by $\mathbf{A}_t = \{A_{1,t}, \ldots, A_{K,t}\} \in \{0,1\}^K$, where, for $k = 1, \ldots, K$, $A_{k,t} = 1$ if treatment k is received and 0 if not at time t, and the corresponding interventional features by $\mathbf{V}_t = (V_{1,t}, \ldots, V_{K,t}) \in \mathcal{V} \subset \mathbb{R}^{D_v \times K}$, where V_k denotes the features of treatment k. In particular, $K > 2$ means the outcome is caused by multiple interventions. It is worth mentioning that, instead of using one-hot treatments of the mixed interventions to correspond to one intervention plan, we explore each of the intervention actions that make up the mixed intervention regimen.

Following the potential outcomes framework proposed by Rubin in 1978 Rubin (1978) and extended in 2008 Robins & Hernan (2008) to account for time-varying treatments, we let $Y[\mathbf{a}_t]$ denote the individual potential outcome at time t for the treatment $\mathbf{A}_t = \mathbf{a}_t$, and $Y_t = Y[\mathbf{A}_t]$ is the individual observed outcome at time t of treatment $\mathbf{A}_t$. Consider a random sample of size N, for each individual $i = 1, \ldots, N$, we observe discrete time series data $\{\mathbf{V}_s^{(i)}, \mathbf{A}_s^{(i)}, \mathbf{X}_s^{(i)}, Y_s^{(i)}\}_{s=1}^t$ for some positive integer t, distributed as $\{\mathbf{V}_s, \mathbf{A}_s, \mathbf{X}_s, Y_s\}_{s=1}^t$.

To adapt the unconfoundedness assumption, we let $\overleftarrow{\mathbf{H}}_t = (\overleftarrow{A}_{t-1}, \overleftarrow{V}_t, \overleftarrow{X}_t)$, where $\overleftarrow{A}_{t-1} = (\mathbf{A}_1, \ldots, \mathbf{A}_{t-1})$ is the historical treatment assignments, $\overleftarrow{V}_t = (V_1, ..., V_t)$ is the adopted interventional feature up to time t, and $\overleftarrow{X} = (X_1, ..., X_t)$ is the time-varying state vector.

Assumption 1. (unconfoundedness) Base on given notations, unconfoundedness assumption at time t is defined as:

$$\{Y[\mathbf{a}_t]\}_{\mathbf{a}_t \in \{0,1\}^K} \perp\!\!\!\perp A_t \mid \overleftarrow{\mathbf{H}}_t \,. \tag{1}$$

[1]Complete background on treatment effect inference is provided in appendix due to the length limit.

Our goal is to estimate the conditional average treatment effect (CATE) for each treatment k, $k = 1, \ldots, K$, in the future. For notational simplicity in defining our CATE, we let $Y[a_{k,t}]$ be the potential outcome $Y[\mathbf{a}_t]$ with treatment k being $a_{k,t} \in \{0, 1\}$ and $a_{j,t} = 0$ for all $j \neq k$, and let $Y[a_{0,t}]$ be the potential outcome with none of the K treatments (i.e. $a_{1,t} = \ldots = a_{K,t} = 0$). Then, the CATE we aim to estimate is defined as

$$\delta[a_{k,t+\tau}] = \mathbb{E}(Y[a_{k,t+\tau}] \mid \overleftarrow{\mathbf{H}}_t) - \mathbb{E}(Y[a_{0,t+\tau}] \mid \overleftarrow{\mathbf{H}}_t)\,, \tag{2}$$

for $a_{k,t+\tau} = 1$, $k = 1, \ldots, K$, $\tau \geq 0$ an integer, where $t + \tau$ represents a future time step. We hope the future individual potential outcome could be estimated when we decide on specific treatment $a_{k,t+\tau}$, to facilitate decision making. For this purpose, we introduce the following assumption.

Assumption 2. (causal interactions) Assuming that the potential outcomes of multiple treatments can be divided into separated treatment effects and causal interactions. That is, we can define the causal interaction for any combination of the K treatments $\mathbf{a}_t \in \{0, 1\}^K$, $\delta_{CI}[\mathbf{a}_t]$, as follows:

$$\begin{aligned}\delta_{CI}[\mathbf{a}_t] = \mathbb{E}((Y[\mathbf{a}_t] - Y[a_{0,t}]) \mid \overleftarrow{\mathbf{H}}_t) \\ - \sum_{k=1}^{K} \mathbb{E}((Y[a_{k,t}] - Y[a_{0,t}]) \mid \overleftarrow{\mathbf{H}}_t)\,.\end{aligned} \tag{3}$$

3 METHOD

Most existing methods estimate individual treatment effects in static setting without considering the time-varying confounders. Those temporal causal inference methods focus on single treatment regimen, ignoring realistic situations of multiple treatment assignments. More importantly, the causal interactions under mixed treatments is difficult to be accurately estimated. Considering the challenges of time-varying bias, multiple treatment selection bias, and causal interactions in time series prediction task, in this section, we propose TCFimt to forecast the future target value under mixed treatments and quantify the individual treatment effects for decision making.

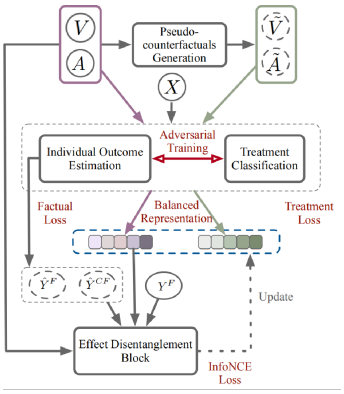


Figure 2: The architecture of our proposed method TCFimt.

The whole framework could be divided into three main parts, as shown in Figure 2. To solve the unobserved counterfactual data problem, we first design a corruption function to generate pseudo-counterfactual data with treatment actions and interventional features. The corruption function design also plays an auxiliary role in subsequent intervention effects separation. Second, we solve the time-varying bias issue and selection bias issue by conducting domain adversarial training in the individual potential outcome estimation task and the multiple treatment action classification tasks, then the balanced representation which is invariant to the treatment at present moment can be learned. In particular, individual outcomes are estimated through an encoder-decoder framework, whose decoder module supports the multi-horizon prediction. Note that due to the similar inputs and shared parameters, only employing the individual outcome estimate module cannot well decouple the colliding intervention effects. We propose to perform balanced representation learning and prediction on both factual and pseudo-counterfactual data. By applying the contrastive learning of factual and pseudo-counterfactual outcome and interaction effect assumption, our method decouple

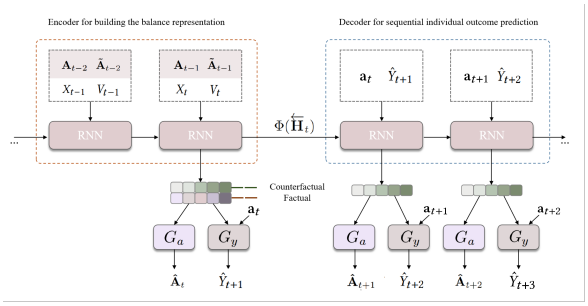


Figure 3: Encoder-decoder framework for potential outcome estimation.

the mixed treatment effect effectively. In the later section, we theoretically show that such a strategy improves estimation accuracy under the setting of multiple treatments. In the third part, an effect disentanglement block is applied. By modeling the relationship between treatments and corresponding potential outcomes via the contrastive learning strategy, this block further highlights the effects of current interventions in the training process.

3.1 Pseudo-counterfactuals Generation

Due to the unobserved nature of counterfactual samples, it brings difficult to train a model with good generalization performance. Especially in the process of time series forecasting task, only one decision is usually adopted at the present moment, while the results of other decisions are unknown. To explore the outcomes under other treatment options and the reasons for not taking them, it is necessary to complete the datasets at first. Instead of learning a joint distribution of causal variables and treatments, we generate the pseudo-counterfactual data with the interventional features $\tilde{V}$ and treatment actions $\tilde{A}$. In detail, a corruption function $\mathbb{C}(\cdot)$ is proposed with $\mathbf{V}$ and $\mathbf{A}$ as input:

$$(\tilde{\mathbf{V}}, \tilde{\mathbf{A}}) = \mathbb{C}(\mathbf{V}, \mathbf{A})\,. \tag{4}$$

The corruption function is implemented based on experiment results, as described below. For each time step a per sample, we randomly specify an intervention position to be changed, then reverse operation is carried out on the intervention features and intervention actions on this position. The reverse operation can be like subtracting by identity vector, to obtain the corresponding counterfactual samples. Our experimental results prove that this strategy is effective.

3.2 Individual Outcome Estimation

The observational data can be used to train a supervised learning model to forecast the final outcome $\mathbf{Y}_t$. However, in the mixed intervention action scenario, the final result is produced by the combined action of multiple treatments, which is not conducive to discovering the role of each treatments/interventions. While potential outcomes under different treatment are unavailable, furthermore, due to the existence of time-varying confounders in time series, it introduces the time-varying bias which leads the model cannot be reliably used for making causal predictions. Hence, inspired by adversarial learning, TCFimt removes this bias through domain adversarial training and estimates the individual potential outcomes both on factual and counterfactual samples, for any intended future treatment assignment.

The basic idea of individual outcome estimation is to learn a balanced treatment-invariant representation at each moment, and subsequently apply the balanced representation and the latest intervention information for potential outcome forecasting in future time steps. Here the balanced representation learning process is most critical.

Balanced representation. In time series data, we observed the historical information $\overleftarrow{\mathbf{H}}_t = (\overleftarrow{\mathbf{A}}_{t-1}, \overleftarrow{\mathbf{V}}_t, \overleftarrow{\mathbf{X}}_t)$ at time step t, which the historical time-varying variable $\overleftarrow{\mathbf{X}}_t$, intervention features $\overleftarrow{\mathbf{V}}_t$ and treatment assignment $\overleftarrow{\mathbf{A}}_{t-1}$ all have an impact on the next potential outcomes. Notice that the current treatment information is excluded from the balanced representation learning for better exploring the impact of the current intervention. To remove the temporal dependence of the

treatments, we hope the balanced representation summarizes the past temporal states $\overleftarrow{\mathbf{H}}_t$, but is not predictive of the treatment $\mathbf{A}_t$. Let Φ denote the representation function that maps the history $\overleftarrow{\mathbf{H}_t}$ to a representation space $\mathcal{R}$. To obtain unbiased treatment effects, Φ needs to construct treatment invariant representations for each treatment K such that $\mathbb{P}(\Phi(\overleftarrow{\mathbf{H}}_t)|A_k = 1) = \mathbb{P}(\Phi(\overleftarrow{\mathbf{H}}_t)|A_k = 0)$. To achieve this and to estimate the potential outcomes under a planned sequence of treatments, the domain adversarial training framework combing with a sequence-to-sequence architecture is proposed.

Encoder-decoder framework. This framework mainly contains two tasks, as shown in Figure 3. By separately predicting the current treatment actions $\hat{\mathbf{A}}_t$ and the next individual outcomes under each treatment option, the balanced representation is learned in the hidden space via the recurrent networks.

Specifically, to estimate the future individual outcomes in the encoder module, we adopt the recurrent network (e.g., RNN) to capture the temporal states of observed factual samples as well as the generated pseudo-counterfactual ones. Intuitively, by inputting the treatment actions $\mathbf{A}(\tilde{\mathbf{A}})$, interventional features $\mathbf{V}(\tilde{\mathbf{V}})$ and time-varying variables X into RNN module, the balanced representation is learned for subsequent prediction in two targets. To aid the representation, we define for time step $s \geq t$, $\overleftarrow{\mathbf{\Omega}}_s = (\overleftarrow{\mathbf{H}}_t, \{\hat{\mathbf{A}}_j\}_{j=t}^{s-1}, \{\hat{Y}_j\}_{j=t+1}^{s})$, where $\hat{\mathbf{A}}_j$ is an estimated treatment assignment at time j, and $\{\hat{Y}_j\}_{j=t+1}^{t} = \emptyset$. Let $\hat{Y}[\mathbf{a}_{t+\tau}]$ be the prediction of individual outcome under the mixed intervention on time step $t + \tau$. The function $\mathbb{G}_y$ makes outcome estimation based on the intervention from the last time step, whose parameter θ_y is shared on factual and counterfactual samples.

$$\hat{\mathbf{Y}}[\mathbf{a}_{t+\tau}] = \mathbb{G}_y(\Phi(\overleftarrow{\mathbf{\Omega}}_{t+\tau}), \mathbf{A}_{t+\tau} = \mathbf{a}_{t+\tau}; \theta_y) \tag{5}$$

The final outcome is calculated by summarizing these estimated individual outcomes, then we can use the outcome to back-propagate. In addition to the individual outcome estimation, another task is to classify the current treatment action, which forms a confrontation for learning the treatment invariant representation. Via K classifiers $\mathbb{G}_a = (\mathbb{G}_{a_1}, \mathbb{G}_{a_2}, ..., \mathbb{G}_{a_K})$ with parameter θ_a, the treatment assigned at the current time is determined.

$$\hat{\mathbf{A}}_t = \mathbb{G}_a(\Phi(\overleftarrow{\mathbf{\Omega}}_t); \theta_a)\,. \tag{6}$$

Similarly, the decoder module also contains these two tasks, but it aims to predict the outcomes of unobserved samples for a sequence of future treatments, which is also called counterfactual prediction. To maintain the temporal information, the decoder network uses the balanced representation computed by the encoder to initialize the state of an RNN. Different from the encoder network, the decoder adopts the auto-regressive way to do predictions on account of the unavailable ground-truth outcomes. Specifically, it uses the (predicted) outcomes from the last time step like $\hat{\mathbf{Y}}_{t+1}$ and $\hat{\mathbf{Y}}_{t+2}$ combining with previously known treatment actions and intervention features as inputs.

3.3 Effect Disentanglement Block

As the prediction from the above individual outcome estimation module is based on the balanced representation learned in a variety of intervention scenarios, it has limitations on the estimation of separated treatment effects and causal interactions. Specifically, the optimization procedure for multiple treatment estimation will mingle the causal effect of each treatment. To tackle this challenge, we design an effect disentanglement block to rectify it by enhancing representation.

Thanks to the pseudo-counterfactuals generation module, we have factual samples and pseudo-counterfactual samples. Let $\hat{Y}^F[a_{k,t+1}]$ and $\hat{Y}^{CF}[\tilde{a}_{k,t+1}]$ be the factual and counterfactual individual outcome under the k-th intervention on time step $t + \tau$. The function $\mathbb{G}_y$ makes outcome estimation based on the intervention from the last time step, whose parameter θ_y is shared on factual and counterfactual samples.

$$\begin{aligned} \hat{\mathbf{Y}}^F[a_{k,t+1}] &= \mathbb{G}_y(\Phi(\overleftarrow{\mathbf{H}}_{t+1}), \mathbf{A}_{t+1} = a_k; \theta_y) \\ \hat{\mathbf{Y}}^{CF}[\tilde{a}_{k,t+1}] &= \mathbb{G}_y(\Phi(\overleftarrow{\mathbf{H}}_{t+1}), \tilde{\mathbf{A}}_{t+1} = \tilde{a}_k; \theta_y) \end{aligned} \tag{7}$$

In order to disentangle the effect under multiple interventions and more accurately learn the relationship between the estimated outcomes and the corresponding interventions, we adopt a mutual information-based contrastive learning strategy.

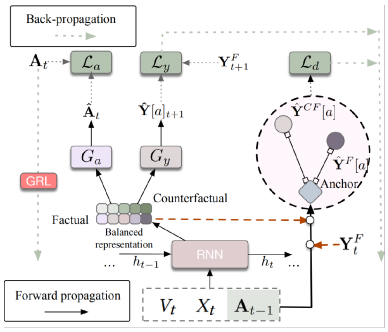


Figure 4: Training Procedure of the whole architecture.

In particular, as shown in the dashed circle of Figure 4, considering the learned balancing representation and the treatment information of factual data, we learn a medium representation and obtain an anchor to distinguish counterfactual from factual data in the estimated potential outcomes. The medium representation $\mathbf{Z}_t = (\mathbf{Z}_{1,t}, ..., \mathbf{Z}_{K,t})$ is calculated via a representation function Ψ with dimension of the number of treatment options and a parameter θ_z:

$$\mathbf{Z}_{k,t} = \Psi(\Phi(\overleftarrow{\mathbf{H}}_t), \mathbf{A}_t; \theta_z)\,. \tag{8}$$

Then the anchor, denoted as $\mathbf{O}_t = (\mathbf{O}_{1,t}, ..., \mathbf{O}_{K,t})$, is calculated with the factual potential outcome $\mathbf{Y}^F$ as:

$$\mathbf{O}_{k,t} = \mathbf{A}_{k,t} \odot \mathbf{Z}_{k,t} \odot \mathbf{Y}^F[a_{k,t}]\,. \tag{9}$$

Based on the anchor and inspired by CPC (contrastive predictive coding)Oord et al. (2018), to measure the relationship on the present state with specific intervention and next individual outcome, we calculate the mutual information $I(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}[a_{k,t+1}])$ between them. For brevity, we denote $\mathbf{O}_{k,t}$ by $\mathbf{O}$ and $\hat{\mathbf{Y}}[a_{k,t+1}]$ by $\hat{\mathbf{Y}}$:

$$I(\mathbf{O}, \hat{\mathbf{Y}}) = \sum_{o,\hat{y}} p(\mathbf{O}, \hat{\mathbf{Y}}) \log \frac{p(\hat{\mathbf{Y}} \mid \mathbf{O})}{p(\hat{\mathbf{Y}})}\,. \tag{10}$$

Similar to the CPC method, we do not predict future potential outcome $\hat{Y}[a_{k,t+1}]$ directly with a generative model $p(\hat{Y}[a_{k,t+1}] \mid O_{k,t})$. Instead, we model a density ratio which preserves the mutual information between $\mathbf{O}_{k,t}$ and $\hat{\mathbf{Y}}[a_{k,t+1}]$ as:

$$f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}[a_{k,t+1}]) \propto \frac{p(\hat{\mathbf{Y}}[a_{k,t+1}] \mid O_{k,t})}{p(\hat{\mathbf{Y}}[a_{k,t+1}])}\,. \tag{11}$$

where $\propto$ means 'proportional to'. The density ratio function is applied as an absolute difference model, to measure the distance between the anchor and individual treatment outcomes:

$$f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}[a_{k,t+1}]) = \mid \mathbf{Z}_{k,t} \odot \hat{\mathbf{Y}}[a_{k,t+1}] - \mathbf{O}_{k,t} \mid\,. \tag{12}$$

After estimating outcomes of multiple treatments $Y[\mathbf{a}_{t+1}]$ and outcomes of individual treatments $Y[a_{k,t+1}]$, the causal interactions $\hat{\delta}_{CI}$ can be estimated for sample i at time t (by Assumption 2):

$$\begin{aligned}\hat{\delta}_{CI} = &(\hat{Y}[\mathbf{a}_{t+1}]^{(i)} - \hat{Y}[a_{0,t+1}]^{(i)}) \\ &- \sum_{k=0}^{K} (\hat{Y}[a_{k,t+1}]^{(i)} - \hat{Y}[a_{0,t+1}]^{(i)})\,.\end{aligned} \tag{13}$$

3.4 Training Procedure

In this subsection, we present the training details of our whole architecture. To remove the bias from time-dependent confounders, it requires that the distance in the distribution of $\Phi(\overleftarrow{\mathbf{H}}_t)$ between any

two pairs of treatments to be minimized. The normally used idea is to assume that the K different treatments represent K pairs of distinct domains (K_{th} treatment implement or not), then based on an adversarial framework of domain adaption to build a representation, which achieves the maximum error in domain classification and the minimum error in outcome prediction.

TCFimt uses domain adversarial training to build a representation of the history $\Phi(\overleftarrow{\mathbf{H}}_t)$ that is both invariant to the treatment action $\mathbf{A}_t$ given at time step t. Specifically, the domain adversarial training is mainly realized through a GRL (Gradient Reversal Layer) Ganin & Lempitsky (2015), which leaves the input unchanged during forwarding propagation and reverses the gradient by multiplying it by a negative scalar during the back-propagation. Further, we add a self-supervised training loss for learning the disentangled intervention effects. The whole training procedure is shown in Figure 4. The total loss consists of three loss functions:

- The treatment classification loss $\mathcal{L}_{a,t}^{(i)}$ is as follows:

$$\mathcal{L}_{a,t}^{(i)} = -\sum_{k=1}^{K} \sum_{j=0,1} \mathbb{I}_{\{a_{k,t}^{(i)}=j\}} \log \left(\mathbb{G}_a^{a_{k,t}=j}(\Phi(\overleftarrow{\mathbf{H}}_t; \theta_r); \theta_a) \right). \tag{14}$$

- The potential outcome forecasting loss $\mathcal{L}_{y,t}^{(i)}$ is as follows:

$$\mathcal{L}_{y,t}^{(i)} = ||\mathbf{Y}_{t+1}^{(i)} - \mathbb{G}_y(\Phi(\overleftarrow{\mathbf{H}}_t; \theta_r); \theta_y)||^2 . \tag{15}$$

- The contrastive learning loss is based on InfoNCE loss:

$$\begin{aligned} \mathcal{L}_{d,t}^{(i)} &= -\sum_{k=0}^{K} \log \frac{f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}^F[a_{k,t+1}])}{f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}^{CF}[a_{k,t+1}])} \\ &= -\sum_{k=0}^{K} \log \frac{\mid \mathbf{Z}_{k,t} \odot \hat{\mathbf{Y}}^F[a_{k,t+1}] - \mathbf{O}_{k,t} \mid}{\mid \mathbf{Z}_{k,t} \odot \hat{\mathbf{Y}}^{CF}[a_{k,t+1}] - \mathbf{O}_{k,t} \mid} . \end{aligned} \tag{16}$$

In a nutshell, our final objective function is to maximize treatment loss, minimize potential outcome loss and minimize the InfoNCE loss Oord et al. (2018). Thus, the overall loss $\mathcal{L}_t^{(i)}$ at timestep t is given by:

$$\mathcal{L}_t^{(i)}(\theta_r, \theta_a, \theta_y, \theta_z) = \sum_{i=1}^{N} \mathcal{L}_{y,t}^{(i)} - \lambda_1 \mathcal{L}_{a,t}^{(i)} + \lambda_2 \mathcal{L}_{d,t}^{(i)} . \tag{17}$$

The hyperparameters λ_1 and λ_2 are used to balance the loss scale. For the choice of the hyperparameters, we start with an initial value for λ_1 and λ_2 use an exponentially increasing schedule during training and get relatively stable results.

3.5 Theoretical Analysis

In this section, we theoretically show that the developed method is effective. Specifically, we show that our overall loss function can effectively bound the expected error of multi-treatment effect estimation. To uncover this, we first show in Theorem 1 that the overall expected error of treatment effect estimation can be bounded by the sum of ϵ_F and ϵ_{CF}, where ϵ_F is the expected factual loss and ϵ_{CF} is the expected counterfactual losses. We note that ϵ_F is upper bounded by $\mathcal{L}_{y,t}^{(i)}$ and transform the inestimable $\epsilon_C F$ (because counterfactual data are not observed) to the difference, ϵ_{CF} - ϵ_F, which is upper bounded by the discrepancy of treated and control distributions. We then show in Theorem 2 that the discrepancy can be upper bounded by $\mathcal{L}_{a,t}^{(i)}$. Therefore, the overall expected error of multi-treatment effect estimation can be bounded by the combination of the potential outcome forecasting loss $\mathcal{L}_{y,t}^{(i)}$ and the treatment classification loss $\mathcal{L}_{a,t}^{(i)}$ in the adversarial framework, which ensures the estimability of multiple treatment effects. Finally, we show in Theorem 3 that minimizing the contrastive learning loss $\mathcal{L}_{d,t}^{(i)}$ maximizes the lower bound of mutual information of anchors and outcomes, which further enhances the balanced representation, facilitating the estimation of separated main treatment effects and causal interactions.

Theorem 1. Upper Bound of Estimation Let $\epsilon_{PEHE}(f)$ denote the expected error in estimating the individual treatment effect of a function[2]. Let $disc(.,.)$ denote the discrepancy, let $p_\Phi^{a_{k,t}}$ be

[2]See the details of definition and proof in the appendix.

treated and control distributions induced by Φ on R, then :

$$\begin{aligned}
&\epsilon_{PEHE}(h,\Phi) \\
&\leq 2(\epsilon_{CF}(h,\Phi) + \epsilon_F(h,\Phi) - 2\sigma_Y^2) \\
&\leq 2(\epsilon_F^{a_{k,t}=0}(h,\Phi) + \epsilon_F^{a_{k,t}=1}(h,\Phi) \\
&+ disc(p_\Phi^{a_{k,t}=1}, p_\Phi^{a_{k,t}=0}))
\end{aligned} \tag{18}$$

According to Theorem 1, ϵ_{PEHE} is upper bounded by the sum of the expected factual loss ϵ_F and expected counterfactual loss ϵ_{CF}. And discrepancy of treated and control distributions can be effective upper bound of the difference $\epsilon_{CF} - \epsilon_F$ of treatment k at time t. The proof of Theorem 1 is based on the theory in CFR Shalit et al. (2017).

Theorem 2. Balanced representation Let P_j denote the distribution of H_t conditional on $a_{k,t} = i$, let $\mathbb{G}_a^{a_k=i}$ denote the output of $\mathbb{G}_a$ for treatment $a_{k,t} = i$. Then the objective for Φ becomes:

$$\begin{aligned}
&\min_{\Phi} \max_{G_a} \sum_{k=1}^{K} \sum_{i=0,1} \mathbb{E}_{\overleftarrow{\mathbf{H}}_t \sim P_{a_k}} [\log(\mathbb{G}_a^{a_k=i}(\Phi(\overleftarrow{\mathbf{H}}_t;\theta_r);\theta_a))] \\
&\qquad s.t. \quad k \in 1,2,..,K, \sum_{i=0,1} \mathbb{G}_a^{a_k=i}(\Phi(\overleftarrow{\mathbf{H}}_t)) = 1 \\
&= \sum_{k=1}^{K} JSD(P_\Phi^{a_k=0}(x') \| P_\Phi^{a_k=1}(x')) - 2Klog2.
\end{aligned} \tag{19}$$

The result of Theorem 2 shows $\mathcal{L}_{a,t}^{(i)}$ has a global minimum which is attained if and only $P_\Phi^{a_k=0} = P_\Phi^{a_k=1}$ for each $k \in 1,2,..,K$. And $\mathcal{L}_{a,t}^{(i)}$ can be replacement of Jensen-Shannon Divergence $JSD(\cdot\|\cdot)$ without the time dependency. Combined with the result of Theorem 1 and Theorem 2, $\mathcal{L}_{a,t}^{(i)}$ can be the upper bound of the difference $\epsilon_{CF} - \epsilon_F$ in temporal multi-intervention.

Theorem 3. Estimating the Mutual Information with InfoNCE Let $I(\cdot,\cdot)$ denote mutual information, then the infoNCE loss:

$$\begin{aligned}
\mathcal{L}_{d,t}^{(i)} &= -\sum_{k=0}^{K} \log \frac{f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}^F[a_{k,t+1}])}{f(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}^{CF}[a_{k,t+1}])} \\
&\geq -\sum_{k=0}^{K} I(\mathbf{O}_{k,t}, \hat{\mathbf{Y}}^F[a_{k,t+1}]).
\end{aligned} \tag{20}$$

Theorem 3 shows the infoNCE loss is a low bound of mutual information between $\mathbf{O}_{k,t}$ and $\hat{\mathbf{Y}}^F[a_{k,t+1}]$. The result proves that $\mathcal{L}_{d,t}^{(i)}$ enhances the representation learning in the estimation of separated treatment outcomes and causal interactions. For details, see the appendix.

4 EXPERIMENTS

We mainly focus on the following research questions:

- Q1: How is the prediction performance of the proposed TCFimt method for future outcomes?
- Q2: What is the auxiliary ability of the TCFimt method for future intervention decision making and the suitable timing of intervention?
- Q3: After the separation of intervention effects, what is the prediction performance of causal interactions between different treatments?

Counterfactual Forecasting Performance. To demonstrate the effectiveness of the TCFimt method on future forecasting for counterfactual samples, we compare it with the above baselines on the Games dataset with two percentage metrics,[3] as shown in Table 1. Since the real intervention are obtained with time-series data in this dataset, the comparison result is more authentic and reliable.

[3]Experimental settings are provided in appendix due to length limit.

Method	Game1		Game2	
	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)
ARMA	6.75	3.88	7.11	3.72
ARIMA	6.52	4.13	6.96	3.92
Prophet	5.91	3.84	6.29	4.11
LSTM	6.20	3.62	7.08	4.36
TCN	5.93	4.69	6.43	5.04
CFR	6.69	4.75	8.04	4.90
TARNet	6.73	4.92	8.02	4.98
CRN	5.94	4.85	6.05	4.31
TCFimt(Ours)	**4.78**	**3.60**	**5.95**	**4.09**

Table 1: Performance of next-one factual prediction on two game datasets.

		$\gamma_c = 5, \gamma_r = 5$			$\gamma_c = 5, \gamma_r = 0$			$\gamma_c = 0, \gamma_r = 5$		
	τ	CRN	TCFimt*	TCFimt	CRN	TCFimt*	TCFimt	CRN	TCFimt*	TCFimt
RMSE	3	6.65%	6.32%	**5.92%**	2.07%	2.03%	**2.02%**	3.09%	2.89%	**2.44%**
	4	7.94%	7.24%	**6.23%**	2.08%	2.01%	**1.96%**	2.46%	2.38%	**2.32%**
	5	8.03%	7.84%	**7.26%**	2.41%	2.32%	**2.26%**	3.06%	2.86%	**2.66%**
	6	7.23%	7.09%	**7.02%**	2.67%	2.63%	**2.42%**	2.76%	2.63%	**2.48%**
	7	9.62%	9.03%	**8.14%**	2.83%	2.73%	**2.58%**	2.81%	2.79%	**2.74%**
Tr Acc.	3	63.0%	72.0%	**74.6%**	69.4%	71.6%	**73.4%**	69.4%	70.4%	**71.4%**
	4	65.2%	70.2%	**73.2%**	71.2%	72.3%	**72.9%**	70.8%	70.8%	**71.0%**
	5	65.4%	69.3%	**72.4%**	70.0%	70.9%	**71.8%**	65.2%	68.2%	**69.4%**
	6	66.2%	68.7%	**71.0%**	65.8%	69.8%	**70.2%**	65.8%	67.3%	**69.3%**
	7	63.4%	65.4%	**70.2%**	65.2%	67.6%	**70.2%**	66.2%	67.2%	**68.8%**
TrT Acc.	3	36.0%	37.3%	**38.8%**	29.6%	37.2%	**38.8%**	35.0%	35.5%	**35.8%**
	4	34.0%	35.3%	**36.0%**	35.9%	36.3%	**36.6%**	34.0%	34.9%	**36.8%**
	5	35.6%	35.6%	**35.8%**	35.8%	36.4%	**37.6%**	31.2%	35.3%	**36.4%**
	6	36.0%	36.4%	**36.8%**	34.0%	35.2%	**36.0%**	33.8%	34.2%	**35.3%**
	7	35.2%	36.3%	**37.2%**	35.2%	36.3%	**36.8%**	39.8%	39.8%	**39.9%**

Table 2: Performance of recommending the correct treatment and timing of treatment on Tumour dataset. $TCFmit^*$ is the TCFmit without effect disentanglement block and contrastive learning.

From the table, we have the following observations: The TCFimt achieves the best performance on the Game1 dataset. For the Game2 dataset, our proposed method also beats all baselines on the RMSE metric and is slightly inferior to the first type of comparison method on the MAE metric. This illustrates that the learned balanced representation by TCFimt is more effective in predicting the potential outcomes of subsequent counterfactual samples. Through the data analysis of Game1 and Game2 datasets, the fluctuations in the Game2 dataset are greater than that of Game1. Therefore, from the comparison of the MAE results on Game2, the ARMA and ARIMA model are more inclined to predict an average result, which leads to the corresponding MAE result is smaller.

Treatment Recommending Performance. To validate the second question Q2 and Q3, we conduct experiments on the Tumour dataset as to which treatment actions were taken and when[4]. We first calculate the RMSE of the prediction of the tumour volume and causal interactions by assumption 2. We then calculate the accuracy of taking the right treatments (Tr Acc.) for patients, and then the accuracy of treatment timing (TrT Acc.) when the right treatment was selected. Based on the such rule, the TrT Acc. is bound to be smaller than Tr Acc. In Table 2, the parameter τ means the predicted horizon, γ_c and γ_r are the hyperparameters to generate distinct Tumour datasets with varied treatment assignments. The results show that our proposed method TCFimt achieves higher accuracy than the CRN method in most cases, i.e., the TCFimt model has certain advantages in assisting the formulation of treatment. The prediction accuracy decrease as the increase of parameter τ, but TCFmit has better performance than the related work at any parameter τ.

5 CONCLUSION

In this paper, we propose a novel method named TCFimt for counterfactual forecasting of time-series data in multiple mixed intervention scenarios. We first design a function to generate pseudo counterfactual samples, which helps to mitigate the selection bias. Via the adversarial loss of potential outcome prediction and treatment classification tasks, a treatment invariant representation was learned for alleviating the time-varying bias. Finally, the intervention mixing problem is solved based on a contrastive learning way, which also enhanced the forecasting of individual potential outcomes. Extensive experiments exhibited the superiority of our proposed model on counterfactual forecasting and choosing the correct treatments. Our study is an attempt at causal inference under mixed treatment effects, it may bring some new insights in this direction.

[4]See the appendix for more details.

REFERENCES

Ahmed M Alaa and Mihaela van der Schaar. Bayesian inference of individualized treatment effects using multi-task gaussian processes. *arXiv preprint arXiv:1704.02801*, 2017.

Ahmed M Alaa, Michael Weisz, and Mihaela Van Der Schaar. Deep counterfactual networks with propensity-dropout. *arXiv preprint arXiv:1706.05966*, 2017.

Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

Ioana Bica, Ahmed M Alaa, James Jordon, and Mihaela van der Schaar. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. *arXiv preprint arXiv:2002.04083*, 2020.

Hugh A Chipman, Edward I George, Robert E McCulloch, et al. Bart: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1):266–298, 2010.

Ping Feng, Xiao-Hua Zhou, Qing-Ming Zou, Ming-Yu Fan, and Xiao-Song Li. Generalized propensity score for estimating the average treatment effect of multiple treatments. *Statistics in medicine*, 31(7):681–697, 2012.

Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pp. 1180–1189. PMLR, 2015.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

Liangyuan Hu, Chenyang Gu, Michael Lopez, Jiayi Ji, and Juan Wisnivesky. Estimation of causal effects of multiple treatments in observational studies with a binary outcome. *Statistical methods in medical research*, 29(11):3218–3234, 2020.

Xia Jiang, Richard E Neapolitan, M Michael Barmada, Shyam Visweswaran, and Gregory F Cooper. A fast algorithm for learning epistatic genomic relationships. In *AMIA annual Symposium proceedings*, volume 2010, pp. 341. American Medical Informatics Association, 2010.

Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pp. 3020–3029, 2016.

Bryan Lim, Ahmed Alaa, and Mihaela van der Schaar. Forecasting treatment responses over time using recurrent marginal structural networks. *NeurIPS*, 18:7483–7493, 2018.

Hao Liu, Yunze Li, Qinyu Cao, Guang Qiu, and Jiming Chen. Estimating individual advertising effect in e-commerce. *arXiv preprint arXiv:1903.04149*, 2019.

Michael J Lopez and Roee Gutman. Estimation of causal effects with multiple treatments: a review and new ideas. *Statistical Science*, pp. 432–454, 2017.

Min Lu, Saad Sadiq, Daniel J Feaster, and Hemant Ishwaran. Estimating individual treatment effect in observational data using random forest methods. *Journal of Computational and Graphical Statistics*, 27(1):209–219, 2018.

Jared K Lunceford and Marie Davidian. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in medicine*, 23(19):2937–2960, 2004.

Saisai Ma, Lin Liu, Jiuyong Li, and Thuc Duy Le. Data-driven discovery of causal interactions. *International Journal of Data Science and Analytics*, 8(3):285–297, 2019.

Ian McLeod et al. Derivation of the theoretical autocovariance function of autoregressive-moving average time series. *Applied Statistics*, 24(2):255–256, 1975.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Kristin E Porter, Susan Gruber, Mark J Van Der Laan, and Jasjeet S Sekhon. The relative performance of targeted maximum likelihood estimators. *The international journal of biostatistics*, 7 (1), 2011.

James M Robins. Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics-Theory and methods*, 23(8):2379–2412, 1994.

James M Robins. Association, causation, and marginal structural models. *Synthese*, pp. 151–179, 1999.

James M Robins and Miguel A Hernan. Estimation of the causal effect of time-varying exposures. In *Longitudinal data analysis*, pp. 547–593. Chapman and Hall/CRC, 2008.

Donald B Rubin. Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, pp. 34–58, 1978.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, pp. 3076–3085. PMLR, 2017.

Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018.

Jinsung Yoon, James Jordon, and Mihaela van der Schaar. Ganite: Estimation of individualized treatment effects using generative adversarial nets. In *International Conference on Learning Representations*, 2018.

G Peter Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175, 2003.

Hao Zou, Peng Cui, Bo Li, Zheyan Shen, Jianxin Ma, Hongxia Yang, and Yue He. Counterfactual prediction for bundle treatment. *Advances in Neural Information Processing Systems*, 33, 2020.

Appendix

In the appendix, we provide additional information about related work, the Rubin-Neyman causal model, the proofs of theorems, and the details of experiments and the case study.

A Related Work

The related work about the future prediction of time series data from the causal inference aspect could be split into three parts, individual treatment effects (ITE) estimation[5], the estimation of treatment effects over time, and the estimation of multiple treatment effects and causal interactions. Previous works on ITE estimation mainly consider the binary treatment situation, which can be divided into three categories. The first one is using a separate model to learn each treatment, such a learning way is biased towards the distribution of different treatment populations, which does not account for the selection bias problem. The second one combines the treatment as a feature, by using one model to learn overall treatments, and the distribution mismatch between the treated and control distributions is adjusted in order to settle selection bias. In this category, non-deep learning methods like the tree-based models Chipman et al. (2010); Lu et al. (2018), doubly-robust methods Porter et al. (2011), propensity and matching based methods Lunceford & Davidian (2004); Lopez & Gutman (2017) are previously proposed. In recent years, some researchers are more focused on using the deep learning approaches in the ITE estimation. For example, Johansson, Shalit, and Sontag (2016) designed a deep representation network to embed the original contexts, to guarantee that the distribution of the contexts after the representation is similar between two different treatments, as well as the small regression loss Johansson et al. (2016). In their later version, a theoretical upper bound on the expected ITE is given to yield a more rigorous estimation algorithm Shalit et al. (2017). However, the methods of balanced representation based on discrepancy measures are not generalize to multiple treatments in temporal settings. The third one is based on the multi-task approach. For example, Alaa et al. used a deep multitask network with a set of shared layers among the factual and counterfactual outcomes. Besides, a multi-task Gaussian process was adopted for bayesian inference of ITE problem Alaa & van der Schaar (2017). Recently, a few works pay more attention to varied forms of treatments that appeared in many real scenarios. For instance, Liu et al.explored the individual effects in the advertising area with multiple continuous treatments Liu et al. (2019), Yoon et al. introduced the idea of a generative adversarial network (GAN) into ITE estimation for fitting any number of treatments Yoon et al. (2018). Moreover, the recent work Zou et al. (2020) learned the latent representation of bundle treatments via variational autoencoder (VAE). Different from these studies, we explore multiple treatments from the new perspective of effect decoupling.

The earliest research on exploring treatment effects over time can be traced back to the Structural Nested Models and Marginal Structural Models (MSMs) Robins (1994; 1999). These methods used predictors performing logistic/linear regression which limited the ability for handling complex time dependencies. To address that, methods based on Bayesian non-parametric or recurrent neural networks as part of these frameworks have been proposed Lim et al. (2018). Moreover, Bica et al. employed an encoder-decoder framework via adversarial framework(CRN) to estimate the individual causal effect of different treatments. Different from CRN, the TCFimt estimates the multiple treatment effects and the combination of treatment effects in temporal via adversarial training (the new loss functions to alleviate selection bias of mixed treatments) and contrastive learning.

Most of the research about multiple treatments and causal interactions focused on average treatment effect(ATE) in the static setting. Multiple treatment methods for ATE include Bayesian Additive Regression Trees (BART) Hu et al. (2020), regression adjustment on the multivariate spline of generalized propensity scores (RAMS) and inverse probability of treatment weighting (IPTW) Feng et al. (2012). And some search algorithms have been developed for causal interaction detection like the Multiple beam search algorithm (MBS) Jiang et al. (2010) and the data-driven approach to causal interaction discovery (DACID) Ma et al. (2019). Different from the above methods, our method provides a view of multiple treatments and causal interactions of individual temporal intervention.

[5]To briefly clarify the usage of ITE (in the appendix), here ITE is actually the estimation of causal effect for each sample of CATE or ATE (mentioned in the main article).

B RUBIN-NEYMAN CAUSAL MODEL

The Rubin-Neyman causal model also called the Rubin causal model (RCM), is a causal inference method based on the framework of the potential outcome. There are several potential outcomes for a few intervention states before the interventions are realized. Once the intervention is achieved, only one potential outcome can be observed, which is the factual outcome, and the other potential outcomes are counterfactual outcomes. We follow the Rubin-Neyman potential outcomes framework. Specifically, let $\mathbf{A} = \{A_1, ..., A_K\} \in \{0,1\}^K$ represent the binary random treatment or intervention vector with K treatments that an individual may receive, $Y[\mathbf{a}]$ denote the potential outcome under treatment $\mathbf{A} = \mathbf{a} \in \{0,1\}^K$ and we also observe some features of the individual, denoted by $\mathbf{X}$. The following two assumptions are commonly made in treatment effect inference:

Assumption 1. (Overlap) For all $k \in \{1, ..., K\}$,

$$0 < \mathbb{P}(A_k = 1 \mid \mathbf{X} = x) < 1\,.$$

This assumption ensures that each intervention is likely to occur in any feature value.

Assumption 2. (Unconfoundedness) Conditioned on $\mathbf{X}$, the potential outcome $Y[\mathbf{a}]$ is independent of $\mathbf{A}$ for all $\mathbf{a} \in \{0,1\}^K$, i.e.

$$\{Y[\mathbf{a}]\}_{\mathbf{a}\in\{0,1\}^K} \perp\!\!\!\perp \mathbf{A} \mid \mathbf{X}\,.$$

This assumption also refers to conditional ignorability, which means after controlling the covariate $\mathbf{X}$, the allocation of individual intervention is independent of the potential outcome. In general, the allocation mechanism of classical randomized experiments satisfies this assumption and the allocation function is already known. While for observational studies, the purpose is to identify the unknown allocation mechanism and thus estimate causal effects.

C PROOFS

C.1 UPPER BOUND OF ESTIMATION

This section will give the proof of an upper bound on the expected error in estimating the individual treatment effect for a given representation. The proof of the upper bound is based on the theory in CFR Shalit et al. (2017) and extends the theory to temporal multi-intervention.

Definition 1. (Estimation of Heterogeneous Effect) Let $\epsilon_{PEHE}(f)$ denote the expected error in estimating the individual treatment effect of a function. Let $f : \mathcal{X} \times \{a_k = 0, 1\} \to \mathcal{Y}$ be an hypothesis. Let $p(x)$ be distribution on $\mathcal{X} \times \{a_k = 0, 1\}$, then:

$$\epsilon_{PEHE}(f) = \int_{\mathcal{X}} (\hat{\delta}[a_{k,t}] - \delta[a_{k,t}])^2 p(x)dx \tag{21}$$

Definition 2. (The Expected Treated and Control Losses) Let $\Phi : \mathcal{X} \to \mathcal{R}$ be a representation function. Let $h : \mathcal{R} \times 0, k \to \mathcal{Y}$ be a hypothesis defined over the representation space $\mathcal{R}$. Let Ψ denote the inverse function of Φ, mapping from R to X. Let $L : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$. be a loss function. Let $p^{a_{k,t}}(x) := p(x, a_{k,t}) = p(x|a_{k,t})$ be the distribution of the features x conditioned on treatment k at time t. let $\ell_{h,\Phi}(x,t) = \int_{\mathcal{Y}} L(Y_t, h(\Phi(x), t))p(Y_t|x)dY_t$ be the expected loss for the unit and treatment pair, then:

$$
\begin{aligned}
\epsilon_F(h,\Phi) &= \int_{\mathcal{X}\times\{0,1\}} \ell_{h,\Phi}(x,a_{k,t})p(x,a_{k,t}=1)dxda_{k,t} \\
\epsilon_{CF}(h,\Phi) &= \int_{\mathcal{X}\times\{0,1\}} \ell_{h,\Phi}(x,a_{k,t})p(x,a_{k,t}=0)dxda_{k,t} \\
\epsilon_F^{a_{k,t}=1}(h,\Phi) &= \int_{\mathcal{X}} \ell_{h,\Phi}(x,a_{k,t}=1)p^{a_{k,t}=1}(x)dxda_{k,t} \\
\epsilon_F^{a_{k,t}=0}(h,\Phi) &= \int_{\mathcal{X}} \ell_{h,\Phi}x,a_{k,t}=0)p^{a_{k,t}=0}(x)dxda_{k,t} \\
\epsilon_{CF}^{a_{k,t}=1}(h,\Phi) &= \int_{\mathcal{X}} \ell_{h,\Phi}(x,a_{k,t}=0)p^{a_{k,t}=1}(x)dxda_{k,t} \\
\epsilon_{CF}^{a_{k,t}=0}(h,\Phi) &= \int_{\mathcal{X}} \ell_{h,\Phi}x,a_{k,t}=1)p^{a_{k,t}=0}(x)dxda_{k,t}
\end{aligned}
\tag{22}
$$

Definition 3. (Discrepancy) Let $disc(.,.)$ denote the discrepancy, let $p_\Phi^{a_{k,t}}$ be treated and control distributions induced by Φ on R:

$$
\begin{aligned}
&disc(p_\Phi^{a_{k,t}=1}, p_\Phi^{a_{k,t}=0})) \\
&= \max_{h,h'\in H} \left| \int_{\mathcal{X}} (p_\Phi^{a_{k,t}=0}(r) - p_\Phi^{a_{k,t}=1}(r)) L(h'(x),h(x))dx \right|
\end{aligned}
\tag{23}
$$

Definition 4. (The expected variance) The expected variance of Y_t with respect to a distribution $p(x,a_{k,t})$:

$$
\begin{aligned}
\sigma^2_{Y_t}(p(x,t)) = \int_{\mathcal{X}\times\mathcal{Y}} &(Y_t - m_t(x))^2 \\
&\times p(Y_t|x)p(x,a_{k,t})dY_t dx
\end{aligned}
\tag{24}
$$

and,

$$
\begin{aligned}
\sigma^2_{Y_t} &= min(\sigma^2_{Y_t}(p(x,a_{k,t}=1)), \sigma^2_{Y_t}(p(x,a_{k,t}=0))) \\
\sigma^2_{Y} &= min(\sigma^2_{Y[a_{k,t}=0]}, \sigma^2_{Y[a_{k,t}=1]})
\end{aligned}
\tag{25}
$$

Lemma 1. Let $u_k = p(a_{k,t}=1), u_0 = p(a_{k,t}=0)$ then:

$$
\begin{aligned}
\epsilon_F(h,\Phi) &= u_k \epsilon_F^{a_{k,t}=1}(h,\Phi) + u_0 \epsilon_F^{a_{k,t}=0}(h,\Phi) \\
\epsilon_{CF}(h,\Phi) &= u_0 \epsilon_{CF}^{a_{k,t}=1}(h,\Phi) + u_k \epsilon_{CF}^{a_{k,t}=0}(h,\Phi)
\end{aligned}
\tag{26}
$$

Proof.(of Lemma 1)

$$
\begin{aligned}
\epsilon_F(h,\Phi) &= \int_{\mathcal{X}\times\{0,1\}} \ell_{h,\Phi}(x,a_{k,t})p(x,a_{k,t})dxda_{k,t} \\
&= p(a_{k,t}=1)\int_{\mathcal{X}} \ell_{h,\Phi}(x,1)p^{a_{k,t}=1}(x)dx \\
&+ p(a_{k,t}=0)\int_{\mathcal{X}} \ell_{h,\Phi}(x,1)p^{a_{k,t}=0}(x)dx \\
&= p(a_{k,t}=1)\epsilon_F^{a_{k,t}=1}(h,\Phi) \\
&+ p(a_{k,t}=0)\epsilon_F^{a_{k,t}=0}(h,\Phi) \\
&= u_k \epsilon_F^{a_{k,t}=1}(h,\Phi) + u_0 \epsilon_F^{a_{k,t}=0}(h,\Phi)
\end{aligned}
\tag{27}
$$

Lemma 2.

$$
\begin{aligned}
\epsilon_{CF}(h,\Phi) \leq &u_0 \epsilon_F^{a_{k,t}}(h,\Phi) + u_k \epsilon_F^{a_{0,t}}(h,\Phi) \\
&+ disc(p_\Phi^{a_{k,t}}, p_\Phi^{a_{0,t}})
\end{aligned}
\tag{28}
$$

Proof.(of Lemma 2)

$$
\begin{aligned}
&\epsilon_{CF}(h,\Phi) - [u_0\epsilon_F^{a_{k,t}=1}(h,\Phi) + u_k\epsilon_F^{a_{k,t}=0}(h,\Phi)] \\
&= [u_0\epsilon_{CF}^{a_{k,t}=1}(h,\Phi) + u_k\epsilon_{CF}^{a_{k,t}=0}(h,\Phi)] \\
&- [u_0\epsilon_F^{a_{k,t}=1}(h,\Phi) + u_k\epsilon_F^{a_{k,t}=0}(h,\Phi)] \\
&= [u_0\epsilon_{CF}^{a_{k,t}=1}(h,\Phi) - u_0\epsilon_F^{a_{k,t}=1}(h,\Phi)] \\
&+ [u_k\epsilon_{CF}^{a_{k,t}=0}(h,\Phi) - u_k\epsilon_F^{a_{k,t}=0}(h,\Phi)]
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
&= u_0 \int_{\mathcal{X}} \ell_{h,\Phi}(x,k)(p^{a_{k,t}=0}(x) - p^{a_{k,t}=1}(x))dx \\
&+ u_k \int_{\mathcal{X}} \ell_{h,\Phi}(x,0)(p^{a_{k,t}=1}(x) - p^{a_{k,t}=0}(x))dx \\
&= u_0 \int_{\mathcal{R}} \ell_{h,\Phi}(\Psi(r),k)(p_\Phi^{a_{k,t}=0}(r) - p_\Phi^{a_{k,t}=1}(r))dr \\
&+ u_k \int_{\mathcal{R}} \ell_{h,\Phi}(\Psi(r),0)(p_\Phi^{a_{k,t}=1}(r) - p_\Phi^{a_{k,t}=0}(r))dr \\
&\leq u_0 \max_{h,h' \in H} \left| \int_{\mathcal{R}} (p_\Phi^{a_{k,t}=0}(r) - p_\Phi^{a_{k,t}=1}(r)) L(h'(x), h(x)dx \right| \\
&+ u_k \max_{h,h' \in H} \left| \int_{\mathcal{R}} (p_\Phi^{a_{k,t}=0}(r) - p_\Phi^{a_{k,t}=1}(r)) L(h'(x), h(x)dx \right| \\
&= \max_{h,h' \in H} \left| \int_{\mathcal{R}} (p_\Phi^{a_{k,t}=0}(r) - p_\Phi^{a_{k,t}=1}(r)) L(h'(x), h(x)dx \right| \\
&= disc(p_\Phi^{a_{k,t}=1}, p_\Phi^{a_{k,t}=0}))
\end{aligned} \tag{30}
$$

Theorem 1.

$$
\begin{aligned}
&\epsilon_{PEHE}(h,\Phi) \\
&\leq 2(\epsilon_{CF}(h,\Phi) + \epsilon_F(h,\Phi) - 2\sigma_Y^2) \\
&\leq 2(\epsilon_F^{a_{k,t}=0}(h,\Phi) + \epsilon_F^{a_{k,t}=1}(h,\Phi) \\
&+ disc(p_\Phi^{a_{k,t}=1}, p_\Phi^{a_{k,t}=0}))
\end{aligned} \tag{31}
$$

Proof.(of Theorem 1) let $m_i(x) = \mathbb{E}(Y_t[a_{k,t} = i] \mid \Phi)$ then:

$$
\begin{aligned}
&\epsilon_{PEHE}(h,\Phi) \\
&= \int_{\mathcal{X}} ((f(x,1) - f(x,0)) - (m_1(x) - m_0(x)))^2 p(x)dx \\
&= 2\int_{\mathcal{X}} (f(x,1) - m_1(x))^2 p(x, a_{k,t} = 1) \\
&+ 2\int_{\mathcal{X}} (m_0(x) - f(x,0))^2 p(x, a_{k,t} = 0) \\
&+ 2\int_{\mathcal{X}} (f(x,1) - m_1(x))^2 p(x, a_{k,t} = 0) \\
&+ 2\int_{\mathcal{X}} (m_0(x) - f(x,0))^2 p(x, a_{k,t} = 1) \\
&= 2\int_{\mathcal{X}} (f(x, a_{k,t} = 1) - m_t(x))^2 p(x, a_{k,t} = 1)dxda_{k,t} \\
&+ 2\int_{\mathcal{X}} (f(x, a_{k,t} = 1) - m_t(x))^2 p(x, a_{k,t} = 0)dxda_{k,t} \\
&\leq 2(\epsilon_{CF}(h,\Phi) + \epsilon_F(h,\Phi) - 2\sigma_Y^2)
\end{aligned} \tag{32}
$$

And according to Lemma 1 and Lemma 2:

$$\begin{aligned} &2(\epsilon_{CF}(h,\Phi)+\epsilon_F(h,\Phi)-2\sigma_Y^2) \\ &\leq 2(\epsilon_F^{a_{k,t}=0}(h,\Phi)+\epsilon_F^{a_{k,t}=1}(h,\Phi) \\ &\quad + disc(p_\Phi^{a_{k,t}=1}, p_\Phi^{a_{k,t}=0})) \end{aligned} \tag{33}$$

C.2 BALANCED REPRESENTATION

The loss function $\mathcal{L}_{a,t}^{(i)}$ has a global minimum which is attained if and only of when the learned representations are invariant across all treatments.

Lemma 3. For fixed Φ, let $x' = \overleftarrow{\mathbf{H}}_t$. The optimal prediction probabilities of G_a are given by:

$$G_a^{a_{k,t}^*}(x') = \frac{P_\Phi^{a_{k,t}}(x')}{\sum_{i=0,1} P_\Phi^{a_{k,t}=i}(x')} \tag{34}$$

Proof.(of Lemma 3) By applying Lagrange multiplies,

$$\begin{aligned} G_a^* = \underset{G_a}{argmax} \sum_{i=0,1} \int_{x'} log(G_a^{a_{k,t}=i}(x')) P_\Phi^{a_{k,t}=i}(x') dx' \\ subject\, to \sum_{i=0,1} G_a^{a_{k,t}=i}(x') = 1 \end{aligned} \tag{35}$$

$$\begin{aligned} G_a^* = \underset{G_a}{argmax} \sum_{i=0,1} log(G_a^{a_{k,t}=i}(x')) P_\Phi^{a_{k,t}=i}(x') \\ +\lambda(\sum_{i=0,1} G_a^{a_{k,t}=i}(x') - 1) \end{aligned} \tag{36}$$

Derive this expression and make it equal to 0

$$\begin{aligned} G_a^{a_{k,t}^*}(x') &= -\frac{P_\Phi^{a_{k,t}}(x')}{\lambda} \\ \lambda &= -\sum_{i=0}^{K} P_\Phi^{a_{k,t}=i}(x') \end{aligned} \tag{37}$$

Theorem 2.

$$\begin{aligned} \min_\Phi \max_{G_a} \sum_{k=1}^{K} \sum_{i=0,1} \mathbb{E}_{\overleftarrow{\mathbf{H}}_t \sim P_{a_{k,t}}} [\log(\mathbb{G}_a^{a_{k,t}=i}(\Phi(\overleftarrow{\mathbf{H}}_t;\theta_r);\theta_a))] \\ s.t. \quad k \in 1,2,..,K, \sum_{i=0,1} \mathbb{G}_a^{a_{k,t}=i}(\Phi(\overleftarrow{\mathbf{H}}_t)) = 1 \end{aligned} \tag{38}$$

has a global minimum which is attained if and only $P_\Phi^{a_{k,t}=0} = P_\Phi^{a_{k,t}=1}$ for each $k \in 1,2,..,K$

Proof.(of Theorem 2) According to Lemma 3

$$
\begin{aligned}
&\min_{\Phi}\max_{G_a}\sum_{k=1}^{K}\sum_{i=0,1}\mathbb{E}_{\overleftarrow{\mathbf{H}}_t\sim P_{a_{k,t}}}[\log(\mathbb{G}_a^{a_{k,t}=i}(\Phi(\overleftarrow{\mathbf{H}}_t;\theta_r);\theta_a))]\\
&s.t.\quad k\in 1,2,..,K,\sum_{i=0,1}\mathbb{G}_a^{a_{k,t}=i}(\Phi(\overleftarrow{\mathbf{H}}_t))=1\\
&=\min_{\Phi}\sum_{k=1}^{K}\mathbb{E}_{x'\sim P_{a_{k,t}}}[\log(\frac{P_\Phi^{a_{k,t}=1}(x')}{P_\Phi^{a_{k,t}=0}(x')+P_\Phi^{a_{k,t}=1}(x')})\\
&+\log(\frac{P_\Phi^{a_{k,t}=0}(x')}{P_\Phi^{a_{k,t}=0}(x')+P_\Phi^{a_{k,t}=1}(x')})]\\
&=\min_{\Phi}\sum_{k=1}^{K}\mathbb{E}_{x'\sim P_{a_{k,t}}}[\log(\frac{P_\Phi^{a_{k,t}=1}(x')}{\frac{1}{2}(P_\Phi^{a_{k,t}=0}(x')+P_\Phi^{a_{k,t}=1}(x'))})\\
&+\log(\frac{P_\Phi^{a_{k,t}=0}(x')}{\frac{1}{2}(P_\Phi^{a_{k,t}=0}(x')+P_\Phi^{a_{k,t}=1}(x'))})]-2Klog2\\
&=\sum_{k=1}^{K}JSD(P_\Phi^{a_{k,t}=0}(x')\|P_\Phi^{a_{k,t}=1}(x'))-2Klog2
\end{aligned}
\tag{39}
$$

Since $JSD(\cdot\|\cdot)$ is Jensen-Shannon Divergence, JSD is non-negative and zero if and only if distributions are equal. And we have that $P_\Phi^{a_{k,t}=0}=P_\Phi^{a_{k,t}=1}$ for each $k\in 1,2,..,K$

C.3 Estimating the Mutual Information with InfoNCE

By minimizing InfoNCE loss, we can maximize the mutual information between $\mathbf{O}_{k,t}$ and $\hat{\mathbf{Y}}^F[a_{k,t+1}]$.

Definition 5. (Mutual Information) the mutual information expression as follows:

$$
f(O_{k,t},Y[a_{k,t+1}])\propto\frac{p(O_{k,t},Y[a_{k,t+1}])}{p(O_{k,t})p(Y[a_{k,t+1}])}. \tag{40}
$$

Theorem 3. The CPC loss is the low bound of mutual Information.

$$
\begin{aligned}
\mathcal{L}_{d,t}^{(i)}&=-\sum_{k=0}^{K}\log\frac{f(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}])}{f(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}\\
&\geq-\sum_{k=0}^{K}I(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}]).
\end{aligned}
\tag{41}
$$

Proof.(of Theorem 3)

$$
\begin{aligned}
\mathcal{L}_{d,t}^{(i)}&=-\sum_{k=0}^{K}\log\frac{f(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}])}{f(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}\\
&=-\sum_{k=0}^{K}\underset{X}{E}log[\frac{\frac{p(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}])}{p(\mathbf{O}_{k,t})p(\hat{\mathbf{Y}}^F[a_{k,t+1}])}}{\frac{p(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}{p(\mathbf{O}_{k,t})p(\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}}]\\
&=-\sum_{k=0}^{K}\underset{X}{E}(log([\frac{p(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}])}{p(\mathbf{O}_{k,t})p(\hat{\mathbf{Y}}^F[a_{k,t+1}])}]\\
&-\underset{X}{E}[\frac{p(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}{p(\mathbf{O}_{k,t})p(\hat{\mathbf{Y}}^{CF}[a_{k,t+1}])}]))\\
&\approx-\sum_{k=0}^{K}\underset{X}{E}(log[\frac{p(\mathbf{O}_{k,t},\hat{\mathbf{Y}}^F[a_{k,t+1}])}{p(\mathbf{O}_{k,t})p(\hat{\mathbf{Y}}^F[a_{k,t+1}])}]-1)\\
&\geq-\sum_{k=0}^{K}I(\mathbf{O}_{k,t},Y_k)
\end{aligned}
\tag{42}
$$

D EXPERIMENTAL SETTINGS

Datasets: For the sake of fully verifying the above questions, we conduct experiments on real-world datasets from two distinct fields.

Games: The dataset contains time-series gaming indicators in distinct game genres coming from NetEase Games. For each game, there is a corresponding game update announcement data, which can be considered as interventions. According to the type of events in the game update announcement, we can get the intervention actions or event options, as summarized in Table 3. These two datasets are all collected from Jan. 2019 to Nov. 2020.

Tumour: With the help of a state-of-the-art bio-mathematical model, this dataset simulates the combined effects of chemotherapy and radiotherapy in lung cancer patients. And at each timestep, there are four treatment options: no treatment, chemotherapy, radiotherapy, combined chemotherapy and radiotherapy. To validate the effectiveness, we adopt the model was based on related wrok Lim et al. (2018) Bica et al. (2020) to evaluate model[6].

The volume of tumour t days after diagnosis is modeled as follows:

$$V(t+1) = \Bigg(1 + \underbrace{\rho \log(\frac{K}{V(t)})}_{\text{Tumor growth}} - \underbrace{\beta_c C(t)}_{\text{Chemotherapy}} - \underbrace{(\alpha_r d(t) + \beta_r d(t)^2)}_{\text{Radiotherapy}} - \underbrace{\beta_m C(t) d(t)}_{\text{Interaction}} + \underbrace{e_t}_{\text{Noise}} \Bigg) V(t)\,. \tag{43}$$

The causal interaction is added by $\beta_m C(t)d(t)$. And for brevity, the other variables in the formula are not explained, the details are the same in the related work Lim et al. (2018) Bica et al. (2020). Here we basically show that the changes in tumour volume are affected by two interventions (i.e., chemotherapy and radiotherapy).

Game ID	#Events	#Event Options	Game Genres
Game1	695	3	action game
Game2	691	2	role-playing game

Table 3: The Statistics of Games Dataset

Baselines: For a fair comparison, we pick baseline methods from three categories. The first group includes classical and widely used time series prediction methods, i.e., ARMA, ARIMA, and Prophet. The second one is the deep neural network related sequential forecasting methods, i.e., LSTM and TCN. From the individual treatment effect estimation perspective, the last type of methods are the latest research approach in this area, like TARNet, CFR, and CRN.

- **ARMA** McLeod et al. (1975) is the autoregressive moving average method.
- **ARIMA** Zhang (2003) is the autoregressive integrated moving average method.
- **Prophet** Taylor & Letham (2018) was proposed by Facebook, is based on decomposition way.
- **LSTM** Hochreiter & Schmidhuber (1997) is a basic recurrent neural network method for sequential prediction.
- **TCN** Bai et al. (2018) combines the convolutional and recurrent networks for sequential modeling.
- **TARNet**[7] Johansson et al. (2016) is the treatment-agnostic representation network, to learn a representation that reduce the discrepancy between the treated and control populations.
- **CFR**[8] Shalit et al. (2017) is the counterfactual regression network, which is similar to TARNet but adding an integral probability metric for bounding the couterfactual loss.
- **CRN**[9] Bica et al. (2020) is the latest counterfactual regression network.

[6]The details of Tumour model in the appendix

[7]https://github.com/clinicalml/cfrnet

[8]https://github.com/clinicalml/cfrnet

[9]https://github.com/ioanabica/Counterfactual-Recurrent-Network/

Evaluation Metrics: For evaluating the forecasting performance in time series, we adopt the normally used RMSE and MAE metrics. As the counterfactual is not available in the games dataset, we treat the testing set in the future time steps as counterfactual samples. Considering question Q2 and Q3, the external accuracy of recommending specific treatment and treatment timing is adopted.

Parameter Settings: Our TCFimt model is implemented with Tensorflow and optimizes the final objective function (in Eq. 17) via the Adam algorithm. The hyperparameters like batch size, learning rate, hidden units in the RNN module, the dropout probability, etc. are searching from a range of values. Details of our implementations can be found in the appendix.

E TREATMENT RECOMMENDING EXPERIMENTS

To validate how well the models help to select the correct treatment and its timing, we implement the treatment recommending experiments on Tumour data with two accuracy metrics. The first one is treatment accuracy, which measures the accuracy of taking the right treatments for patients. The second one is treatment timing accuracy, which is to calculate the degree of time matching under the premise of the correct choice of treatment.

According to Eq. 43, the test data is generated with 500 samples. Let τ denote the future time horizon, $\overleftarrow{\mathbf{H}}$ be the current history of patients. Similar to the construction in related work Bica et al. (2020), we generate 3τ counterfactual data with treatment options: no treatment (A_0), chemotherapy (A_1), radiotherapy (A_2) and both chemotherapy and radiotherapy (A_3).

chemotherapy application:

$$\begin{aligned}
&\mathbf{Y}_{t+\tau} \mid a_t = A_1, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_1, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad ... \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_1, \overleftarrow{\mathbf{H}}
\end{aligned} \tag{44}$$

radiotherapy application:

$$\begin{aligned}
&\mathbf{Y}_{t+\tau} \mid a_t = A_2, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_2, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad ... \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_2, \overleftarrow{\mathbf{H}}
\end{aligned} \tag{45}$$

chemotherapy and radiotherapy application:

$$\begin{aligned}
&\mathbf{Y}_{t+\tau} \mid a_t = A_3, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_3, ..., a_{t+\tau-1} = A_0, \overleftarrow{\mathbf{H}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad ... \\
&\mathbf{Y}_{t+\tau} \mid a_t = A_0, a_{t+1} = A_0, ..., a_{t+\tau-1} = A_3, \overleftarrow{\mathbf{H}}
\end{aligned} \tag{46}$$

For each patient, the number of counterfactuals is $3 \cdot \tau \cdot$ time steps. Here the time steps are set as 60 days.

Hyperparameter	Search range for encoder
Iterations of Hyperparameter Search	200
Learning rate	0.1, 0.01, 0.001, 0.0001
Minibatch size	64, 128, 256, 512
RNN hidden units	0.5C, 1C, 2C, 3C, 4C
Representation size	0.5C, 1C, 2C, 3C, 4C
FC hidden units	0.5R, 1R, 2R, 3R, 4R
RNN dropout probability	0.3, 0.4, 0.5, 0.6

Table 4: The Hyperparameter Search Range for TCFimt Encoder Model.

Hyperparameter	Search range for decoder
Iterations of Hyperparameter Search	200
Learning rate	0.1, 0.01, 0.001, 0.0001
Minibatch size	64, 128, 256, 512
RNN hidden units	Representation size of encoder
Representation size	0.5C, 1C, 2C, 3C, 4C
FC hidden units	0.RC, 1R, 2R, 3R, 4R
RNN dropout probability	0.3, 0.4, 0.5, 0.6

Table 5: The Hyperparameter Search Range for TCFimt Decoder Model

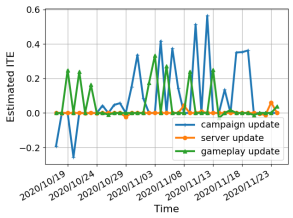
The treatment choosing rule is as follows: since the smaller the tumour volume, the better the treatment effect. According to the estimated outcomes of τ time steps, we select the lowest one among them. Then the correct treatment can be compared, in the case of choosing the correct treatment, we further verify the accuracy of the treatment timing.

Here we provide the hyperparameter search range for the hyperparameter search range for TCFimt in Table 4 and Table 5. And C is the size of the input and R is the size of the balancing representation.

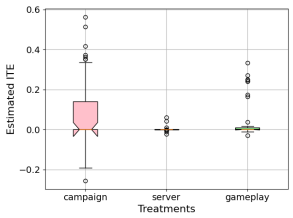
F CASE STUDY

For satisfying the aid of the results to the subsequent decisions, we analyze the individual effects per treatment based on the factual and counterfactual individual estimation outcomes. Please see the case study experiments for more details, here we show the individual effects under the two-game datasets as shown in Figure 5. The left figures are the value of estimated ITE overtime on two games, the right figures are the boxplots of the estimated ITE value under distinct treatments.

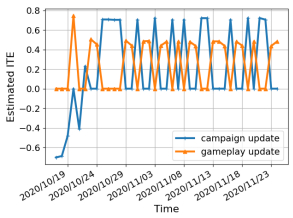
The positive values in the left figures indicate that the corresponding treatment option had a positive effect, and vice versa. As for the Game1 dataset, there are three treatment options, i.e., campaign update, server update, and gameplay update. From the results in Figure 5(a), the first and third treatment options produced larger intervention responses at certain points, while the second treatment option had little or no effects. In combination with Figure 5(b), there are many abnormal points in the figure, which corresponds to the existence of some extreme treatment effect points in the left figure. Comparing the results of the three treatments, the campaign update basically produces positive effects, but its effect fluctuates greatly. The gameplay update in most cases plays a relatively small effect, but in some moments it can also produce positive effects.



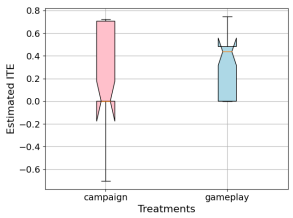
(a) Estimated ITE over time on Game1



(b) Boxplot of Estimated ITE on Game1



(c) Estimated ITE over time on Game2



(d) Boxplot of Estimated ITE on Game2

Figure 5: The estimated ITE results of Game dataset

For the Game2 dataset with two treatment options (i.e., campaign update and gameplay update), the situation seems different. As can be seen from Figure 5(c), the effects of the two treatments appear alternately. From Figure 5(d), the average effects of the campaign update are around zero, but its positive fluctuation is larger than that of the gameplay update, indicating that the good game campaign setting can play a greater role. Besides, the average effect of the second treatment is much greater than that of the first, indicating the importance of the gameplay update in the game. At the same time, the bottom half of the light blue boxplot shows that if the gameplay is poorly designed, it will not have much of an effect.

G CASE STUDY EXPERIMENTS

Due to the unavailable of true treatment effects under each treatment. In the case study, we analyze the ITE based on the estimated individual outcomes of factual and counterfactual data. In detail, the estimated ITE value for each sample i of each treatment k at time t in Figure 5 is calculated as follows:

$$\hat{ITE}_{k,t}^{(i)} = \hat{Y}^{(i)}[a_{k,t}] - \hat{Y}^{(i)}[a_{0,t}] \tag{47}$$

Here we put the factual samples and counterfactual samples together. Recall that in the corruption function, each counterfactual sample at time step t is generated from the factual sample at the same time. Therefore, we always can find the treated and controlled groups for Eq. 47. The left part of Figure 5 can be obtained directly according to the estimated ITE, while we draw the right part by removing the time dimension of the value and keeping the treatment dimension.

The Overview of Treatments in Game Dataset. In our Game dataset, there are three kinds of in-game events, which are also called treatments:

- The campaign update is more like a marketing strategy, which provides special offers or some package sales to promote the consumption or reception of some in-game items (like weapons or skins).
- The server update is an important operation measure for game ecology, which adds new servers for increasing newcomers or merges the declining servers for maintaining the experiences of the active players.
- The gameplay update is a routine adjustment for online games, which calibrates some settings of gameplays or rules for improving the game experience and repairing some found defects.