Hindawi
Journal of Electrical and Computer Engineering
Volume 2017, Article ID 1048385, 11 pages
https://doi.org/10.1155/2017/1048385

Research Article

An Online Causal Inference Framework for Modeling and Designing Systems Involving User Preferences: A State-Space Approach

Ibrahim Delibalta,[1] Lemi Baruh,[1] and Suleyman Serdar Kozat[2]

[1]*Koc University, Istanbul, Turkey*
[2]*Bilkent University, Ankara, Turkey*

Correspondence should be addressed to Ibrahim Delibalta; idelibalta13@ku.edu.tr

Received 2 March 2017; Revised 20 April 2017; Accepted 3 May 2017; Published 22 June 2017

Academic Editor: Zhixin Yang

Copyright © 2017 Ibrahim Delibalta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We provide a causal inference framework to model the effects of machine learning algorithms on user preferences. We then use this mathematical model to prove that the overall system can be tuned to alter those preferences in a desired manner. A user can be an online shopper or a social media user, exposed to digital interventions produced by machine learning algorithms. A user preference can be anything from inclination towards a product to a political party affiliation. Our framework uses a state-space model to represent user preferences as latent system parameters which can only be observed indirectly via online user actions such as a purchase activity or social media status updates, shares, blogs, or tweets. Based on these observations, machine learning algorithms produce digital interventions such as targeted advertisements or tweets. We model the effects of these interventions through a causal feedback loop, which alters the corresponding preferences of the user. We then introduce algorithms in order to estimate and later tune the user preferences to a particular desired form. We demonstrate the effectiveness of our algorithms through experiments in different scenarios.

1. Introduction

Recent innovations in communication technologies, coupled with the increased use of Internet and smartphones, greatly enhanced institutions' ability to gather and process an enormous amount of information on individual users on social networks or consumers in different platforms [1–4]. Today, many sources of information from shares on social networks to blogs, from intelligent device activities to security camera recordings are easily collectable. Efficient and effective processing of this "big data" can significantly improve the quality of many real life applications or products, since this data can be used to accurately profile and then target particular users [5–7]. In this sense, abundance of new sources of information and previously unimaginable ways of access to consumer data have the potential to substantially change the classical machine learning approaches that are tailored to extract information with rather limited access to data using relatively complex algorithms [8–11].

Furthermore, unlike applications where the machine learning algorithms are used as mere tools for processing and inferring using the available data such as predicting the best movie for a particular user [12], the new generation of machine learning systems employed by enormously large and powerful data companies and institutions have the potential to change the underlying problem framework, that is, the user itself, by design [8, 13]. Consider the Google search engine platform and its effects on user preferences. The Google search platform not only provides the most relevant search results but also gathers information on users and provides well-tuned and targeted content (from carefully selected advertisements to specifically selected news) that may be used to change user behavior, inclinations, or preferences [14].

Online users are exposed to persuasive technologies and are continually immersed in digital content and interventions in various forms such as advertisements, news feeds, and recommendations [15]. User decisions and preferences are affected by these interventions [16]. We define a feedback

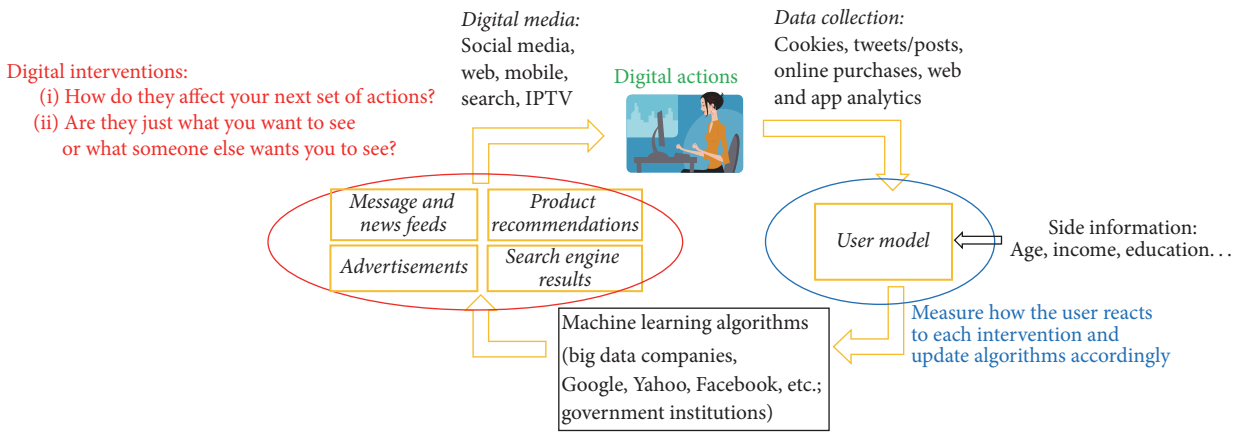


Figure 1: The Digital Feedback Loop.

framework in which these interventions can be selected in a systematic way to steer users in a desired manner. In Figure 1, we introduce "The Digital Feedback Loop" on which we base our model.

To this end, in this paper, we are particularly interested in the causal effects of machine learning algorithms on users [17, 18]. Specifically, we introduce causal feedback loops to accurately describe effects of machine learning algorithms on users in order to design more functional and effective machine learning systems [18, 19]. We model the latent preferences and/or inclinations of a user, as an unknown state in a real life causal system, and build novel algorithms to estimate and, then, alter this underlying unobservable state in an intentional and preferred manner. In particular, we model the underlying evolution of this state using a state-space model, where the latent state is only observed through the behavior of the user such as his/her tweets and Facebook status shares. The internal state is causally affected by the outputs of the algorithm (or the actions of the company), which can be derived from the past observations on the user or outputs of the system. The purpose of the machine learning algorithm can be, for example, (i) to drive the internal system state towards a desired final state, for example, trying to change the opinion of the population towards a newly introduced product; (ii) to maximize some utility function associated with the system, for example, enticing the users to a new and more profitable product; or (iii) to minimize some regret associated with the disclosed information, for example, minimizing the effects of unknown system parameters. Alternatively, the machine learning system may try to achieve a combination of these objectives.

This problem framework readily models a wide range of real life applications and scenarios [18, 19]. As an example, an advertiser may aim to direct the preferences of his/her target audience towards a desired product, by designing advertisement using data collected by consumer behavior surveys [18]. This framework is substantially different from the classical problem of targeted advertisement based on user profiling. In the case of targeted advertising, the goal is to match the best advertisement to the current user, based on the user's profile. Another part of the classical problem is to measure the true impact of an ad (a "treatment" or an "intervention" in the general case) and thus find its effectiveness to help the ad selection for the next time or the next user as well as for billing purposes. Here, we assume that the underlying state, that is, the preferences of the consumers, are not only used to recommend a particular product but also intentionally altered by our algorithm. As in some of the earlier works [12, 17, 20], we use a causal framework to do our modeling. We then take it a step further to mathematically prove that the impact of a treatment can be predesigned and the user can, in theory, be swayed in accordance with the designer's intent. To the best of our knowledge, this is unique to our work. We can further articulate the difference between our work and some of the earlier works using an example in the context of news recommendation. The classical approach tries to show the user news articles he/she might be interested in reading, based on their profile and possibly some other contextual data. A separate process collects information on whether the user clicked on a particular news item and what that item's context is. This collected data is then used to augment the user's profile so that the recommendation part of the process makes a better decision the next time or for the next user. The connection between separate decisions is mainly the enhanced user profile. In reality, the recommended news articles have impacted the user's news preferences to some degree. This is a classical counterfactual problem [8]. While the user preferences themselves are latent and cannot be directly measured, the impact manifests itself in a number of ways that are observable. For instance, the user might tweet about that news with a particular sentiment or buy a book online which is related to the topic in the news item. What we prove with our framework is that, using the observable data and our model, one can produce a sequence of actions which will influence and steer the user's preferences in a pattern that is intended by the recommender system. These actions can be in the form of content served to the user such as news articles, social media feeds, and search results.

In different applications the preferences can be the state and the advertisements (content, the medium of the

advertisement, the frequency, etc.) are the actions or output of the machine learning algorithm. In a different context, the opinions of the social network users on Facebook of a particular event or a new product can be represented as a state. Our model is comprehensive such that the relevant information on the user such as his/her age, gender, demographics, and residency is collectively represented by a side information vector since the advertiser collects data on the consumer such as the spending patterns, demographics, age, gender, and polls.

A summary of our work in this paper is as follows, with the last bullet being our key contribution:

(i) We model the effects of machine learning algorithms such as recommendation engines on users through a causal feedback loop. We introduce a complete state-space formulation modeling: (1) evolution of preferences vectors, (2) observations generated by users, and (3) causal feedback effects of the actions of algorithms on the system. All these parameters are jointly optimized through an Extended Kalman Filtering framework.

(ii) We introduce algorithms to estimate the unknown system parameters with and without feedback. In both cases, all the parameters are estimated jointly. We emphasize that we provide a complete set of equations covering all the possible scenarios.

(iii) To tune the preferences of users towards a desired sequence, we also introduce a linear regression algorithm and introduce an optimization framework using stochastic gradient descent algorithm. Unlike all the previous works that only use the observations to predict certain desired quantities, as the first time in the literature, we specifically design outputs to "update" the internal state of the system in a desired manner.

The rest of the paper is organized as follows. In the next section, we present a comprehensive state-space model that includes the evolution of the latent state vector, underlying observation model and side information. In the same section, we also introduce the causal feedback loop and possible variations to model different real life applications. We then introduce the Extended Kalman Filtering framework to estimate the unknown system parameters. We investigate different real life scenarios including the system with and without the feedback. We present all update and estimation equations. In the following section, we introduce an online learning algorithm to tune the underlying state vector, that is, preferences vector, towards a desired vector sequence through a linear regression and causal feedback loop. We then demonstrate the validity of our introduced algorithms under different scenarios via simulations. We include our simulation results to show that we are able to converge on unknown parameters in designing a system which can steer user preferences. The final section includes conclusions and scope of future work.

2. A Mathematical Model for User Preferences with Causal Feedback Effects

In this paper, all vectors are column vectors and denoted by lower case letters. Matrices are represented by uppercase letters. For a vector $\mathbf{u}$,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}^T\mathbf{u}} \tag{1}$$

is the l^2-norm, where $\mathbf{u}^T$ is the ordinary transpose. For vectors $\mathbf{a} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{a}^T$ is the transpose and $[\mathbf{a}; \mathbf{b}] \in \mathbb{R}^{m+n}$ is the concatenated vector. Here, $\mathbf{I}$ represents an identity matrix, $\mathbf{0}$ represents a vector or a matrix of all zeros, and $\mathbf{1}$ represents a vector or a matrix of all ones, where the size is determined from the context. The time index is given in the subscript; that is, $\mathbf{x}_t$ is the sample at time $\mathbf{t}$. δ_t is the Kronecker delta functions.

We represent preferences of a user as a state vector $\mathbf{p}_t$, where this state vector is latent; that is, its entries are unknown by the system designer. The state vector can represent affinity or opinions of the underlying social network user for different products or for controversial issues like privacy. The actual length and values of the preferences depend on the application and context. As an example for the mood of a person in a context of 6 feelings (happy, excited, angry, scared, tender, and sad), the preference vector might be $[0, 1, 0, 0, 0, 0]^T$.

The relevant information on the user such as his/her age, gender, demographics, and residency is collectively represented by a side information vector $\mathbf{s}_t$. The side information on users on the social networks can be collected based on their profiles or their friendship networks. We assume that the side information is known to the designer and, naturally, change slowly so that $\mathbf{s}_t = \mathbf{s}$ is constant in time.

The machine learning system collects data on the user, say $\mathbf{x}_t$, such as Facebook shares, comments, status updates, and spending patterns, which is a function of his/her preferences $\mathbf{p}_t$ and the side information $\mathbf{s}$, given by

$$\mathbf{x}_t = F_t\left(\mathbf{p}_t, \mathbf{s}\right), \tag{2}$$

where the functional relationship $F(\cdot)$ will be clear in the following. Since the information collection process may be prone to errors or misinformation, for example, untruthful answers in surveys, we extend (2) to include these effects as

$$\mathbf{x}_t = F_t\left(\mathbf{p}_t, \mathbf{s}\right) + \mathbf{n}_t, \tag{3}$$

where $\mathbf{n}_t$ is a noise process independent of $\mathbf{p}_t$ and $\mathbf{s}$. We can use other approaches instead of an additive noise model; however, the additive noise model is found to accurately model unwanted observation noise effects [21]. We use a time varying linear state-space model to facilitate the analysis such that we have

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{p}_t + \mathbf{n}_t, \tag{4}$$

where $\mathbf{F}_t$ is the observation matrix [22] corresponding to the particular user and $\mathbf{n}_t$ is i.i.d. with

$$E\left[\mathbf{n}_t\mathbf{n}_r^T\right] = \delta_{t-r}\mathbf{R}, \tag{5}$$

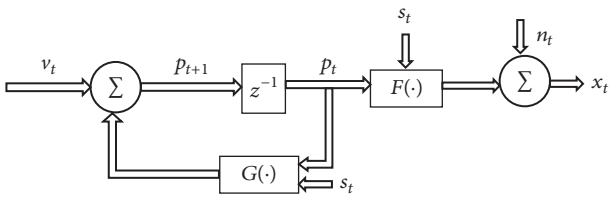


Figure 2: A state-space model to represent evaluation of the user preferences without feedback effects.

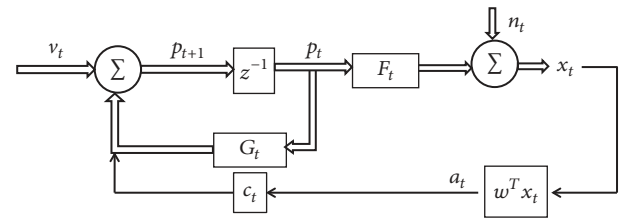


Figure 3: A complete state-space model of the system with action generation and feedback effects.

where $\mathbf{R}$ is the autocorrelation matrix. The autocorrelation matrix $\mathbf{R}$ is assumed to be known, since it can be readily estimated from the data [22] in a straightforward manner. We do not explicitly show the effect of $\mathbf{s}$ on $\mathbf{F}$ for notational simplicity.

Based on prior preferences, different user effects and trends, and the preferences of the user change, we represent this change as

$$\mathbf{p}_{t+1} = G_t\left(\mathbf{p}_t, \mathbf{s}\right) + \mathbf{n}_t, \tag{6}$$

with an appropriate $G_t(\cdot)$ function. To facilitate the analysis, we also use a state-space model

$$\mathbf{p}_{t+1} = \mathbf{G}_t\mathbf{p}_t + \mathbf{v}_t, \tag{7}$$

where $\mathbf{G}_t$ is the state update matrix, which is usually close to an identity matrix since the preferences of user cannot rapidly change [19, 20]. Here, $\mathbf{v}_t$ models the random fluctuations or independent changes in the preferences of users, where it is i.i.d. with

$$E\left[\mathbf{v}_t\mathbf{v}_r^T\right] = \delta_{t-r}\mathbf{Q} \tag{8}$$

and $\mathbf{Q}$ is the autocorrelation matrix. The autocorrelation matrix $\mathbf{Q}$ is assumed to be known, since it can be readily estimated from the data [22] in a straightforward manner. The model without the feedback effects is shown in Figure 2.

Remark 1. To include local trends and seasonality effects, one can use $\mathbf{v}_t = \mathbf{B}_t\mathbf{u}_t$, where $\mathbf{B}_t$ may not be full rank when local trends exist (local trends can cause some data points to be derived from others). Also, $\mathbf{u}_t$ is an i.i.d. noise process. Our derivations in the next sections can be generalized to this case by considering an extended parameter set.

In the following, we model the effect of the actions of the machine learning algorithm in the "observation" (4) and "evolution" (7) equations.

2.1. Causal Inference through the Actions of the Machine Learning System. Based on the collected data $\mathbf{x}_t$, the algorithm takes an action represented by a_t. The action of the machine learning system or the platform can be either discrete or continuous valued depending on the application [21]. As an example, if the action represents a campaign advertisement to be sent to a particular Facebook user, then the set of campaign ads is finite. On the other hand, the action of the machine learning system can be continuous such as providing money incentives to particular users to perform certain tasks such as filling questionnaires. We model the action as a function of the observations as

$$a_t = W_t^T\left(\mathbf{x}_t\right), \tag{9}$$

where $W(\cdot)$ may correspond to different regression methods [21]. To facilitate the analysis, we model the action generation using a linear regression model as

$$a_t = W_t^T\mathbf{x}_t. \tag{10}$$

If we have a finite set of actions, that is, $a_t \in \{1, \ldots, K\}$, we replace (10) by

$$a_t = Q\left(\mathbf{w}_t^T\mathbf{x}_t\right), \tag{11}$$

which is similar to saturation or sigmoid models [23], where $Q(\cdot)$ is an appropriate quantizer. The linear model in (11) can be replaced by more complex models since $\mathbf{x}_t$ can contain discrete entries such as gender and age. However, we can closely approximate any such complex relations by piecewise linear models [24]. The piecewise linear extension of (11) is straightforward [24].

Based on the actions of the machine learning algorithm (and prior preferences), we assume that the preferences of the user changes in a linear state-space form with an additive model for the causal effect [18–20], which yields the following state model:

$$\mathbf{p}_{t+1} = \mathbf{G}_t\mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t a_t, \tag{12}$$

where $\mathbf{c}_t$ is the unknown causal effect. The complete linear state-space model is illustrated in Figure 3. Although there exists other models for the feedback, apart from the linear feedback, the linear feedback was found to accurately model a wide range of real life scenarios provided that causal effects are moderate [19], which is typically the case for social networks; that is, advertisements usually do not have drastic effects on user preferences [19, 20]. Our linear feedback model can be extended to piecewise linear models to approximate smoothly varying nonlinear models in a straightforward manner.

Remark 2. We can also use a jump state model to represent the causal effects for the case where a_t is coming from a finite set. In this case, as an example, the causal effects will change the state behavior of the overall system through a jump state model as

$$\mathbf{p}_{t+1} = \mathbf{G}_{t,Q(\mathbf{w}^T\mathbf{x}_t)}\mathbf{p}_t + \mathbf{v}_t. \tag{13}$$

Our estimation derivations in the following sections can also be extended to cover this case using a jump state model [22].

Remark 3. For certain causal inference problems, the actions sequence a_t may be required to be predictive of some reference sequence d_t, in a traffic prediction context, to sway driver preferences $\mathbf{p}_t$ in a certain direction by disclosing estimates a_t for a certain road d_t, using some publicly available data $\mathbf{x}_t$. To account for these types of scenarios, we complement the model in (12) by introducing

$$d_t = H_t(\mathbf{p}_t) + \boldsymbol{\sigma}_t, \tag{14}$$

where $\boldsymbol{\sigma}_t$ is i.i.d. In this case, the feedback loop will be designed in order to tune d_t to a particular value.

In the following, we introduce algorithms that optimize $\mathbf{w}_t$ so that the overall system behaves in a desired manner given the corresponding mathematical system. However, we emphasize the overall system parameters including the feedback loop parameters are not known and should be estimated only from the available observations $\mathbf{x}_t$. Hence, we carry out both the estimation and design procedures together for a complete system design.

3. Design of the Overall System with Causal Inference

We consider the problem of designing a sequence of actions $\{a_t\}_{t\geq 1}$ in order to influence users based on our observations $\{\mathbf{x}_t\}_{t\geq 1}$, where behavior of the user is governed by his/her hidden preference sequence $\{\mathbf{p}_t\}_{t\geq 1}$. The machine learning system is required to choose the sequence $\{\mathbf{w}_t\}_{t\geq 1}$ in order to accomplish its specific goal. The specific goal naturally depends on the application. As an example, in social networks, the goal can be to change the opinions of users about a new product by sending the most appropriate content such as news articles and/or targeted tweets. In its more general form, we can represent this goal as a utility function and optimize the cumulative gain:

$$\max_{\mathbf{w}_{t\geq 1}} \quad \sum_{t=1}^{\infty} E\left[U_t\right], \tag{15}$$

where $U_t = U_t(\mathbf{p}_t)$ is an appropriate utility function for a specific application. To facilitate the analysis, we choose the utility function as the negative of the squared Euclidean distance between the actual consumer preference $\mathbf{p}_t$ and some desired state $\mathbf{q}_t$. We emphasize that, as shown later in the paper, our optimization framework can be used to optimize any utility function provided that it has continuous first-order derivatives due to the stochastic gradient update. In this case (15) can be written as

$$\min_{\mathbf{w}_{t\geq 1}} \quad \sum_{t=1}^{\infty} E\left[\left\|\mathbf{p}_t - \mathbf{q}_t\right\|^2\right]. \tag{16}$$

The overall system parameters, $\{\mathbf{F}, \mathbf{G}, \mathbf{c}\}$, are not known and should be estimated from our observations. We introduce an Extended Kalman Filtering (EKF) approach to estimate the unknown parameters of the system. We separately consider the estimation framework without the feedback loop, that is, $\mathbf{w} = \mathbf{0}$, and with the feedback loop, that is, $\mathbf{w} \neq \mathbf{0}$. Clearly the estimation task for $\{\mathbf{F}, \mathbf{G}\}$ can be carried out before we produce our suggestions $\mathbf{w}$. In this case, we can estimate these parameters with a better accuracy without the feedback effects since we need to estimate a smaller number of parameters under less complicated noise processes. However, for certain scenarios where this feedback loop is already active, we also introduce a joint estimation framework for all parameters. A system with feedback is more general, realistic, and comprehensive. And feedback is needed in order to tune or influence the preferences of a user in a desired manner. However, a system with feedback is more complex to design and analyze. Therefore, we first provide the analysis for a system without feedback and build on it for an analysis of a system with feedback. After we get the estimated system parameters, we introduce online learning algorithms in order to tune the corresponding system to a particular target internal state sequence, which can be time varying, nonstationary, or even chaotic [23, 25].

3.1. Estimating the Unknown Parameters of the System without Feedback. Without the feedback loop, the system is described by

$$\mathbf{p}_{t+1} = \mathbf{G}_t \mathbf{p}_t + \mathbf{v}_t, \tag{17}$$

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{p}_t + \mathbf{n}_t, \tag{18}$$

where $\mathbf{v}_t$ and $\mathbf{n}_t$ are assumed to be Gaussian with correlation matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively. We then define

$$\boldsymbol{\theta}_t \triangleq \left[\mathbf{G}_t(:); \mathbf{F}_t(:)\right], \tag{19}$$

where $\mathbf{G}_t(:)$ is the vectorized $\mathbf{G}_t$; that is, the columns of $\mathbf{G}_t$ are stacked one after another to get a full column vector. To jointly estimate $\mathbf{p}_t$ and $\boldsymbol{\theta}_t$, we formulate an EKF framework by considering

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t, \tag{20}$$

where $\boldsymbol{\varepsilon}_t$ is the noise in estimating $\boldsymbol{\theta}_t$ through the EKF. Then, using (17) and (20) and considering $\mathbf{p}_t$ and $\boldsymbol{\theta}_t$ as the joint state vector, we get

$$\begin{aligned} \mathbf{x}_t &= f_1(\boldsymbol{\theta}_t, \mathbf{p}_t) + \mathbf{n}_t, \\ \begin{pmatrix} \mathbf{p}_{t+1} \\ \boldsymbol{\theta}_{t+1} \end{pmatrix} &= \begin{pmatrix} f_2(\boldsymbol{\theta}_t, \mathbf{p}_t) \\ \boldsymbol{\theta}_t \end{pmatrix} + \begin{pmatrix} \mathbf{v}_t \\ \boldsymbol{\varepsilon}_t \end{pmatrix}, \end{aligned} \tag{21}$$

where

$$\begin{aligned} f_1(\boldsymbol{\theta}_t, \mathbf{p}_t) &\triangleq \mathbf{F}_t \mathbf{p}_t, \\ f_2(\boldsymbol{\theta}_t, \mathbf{p}_t) &\triangleq \mathbf{G}_t \mathbf{p}_t \end{aligned} \tag{22}$$

are the corresponding nonlinear equations so that we require the EKF framework. The corresponding EKF equations to estimate the augmented states are recursively given as

$$\begin{aligned}
\begin{pmatrix}\mathbf{p}_{t|t}\\ \boldsymbol{\theta}_{t|t}\end{pmatrix} &= \begin{pmatrix}\mathbf{p}_{t|t-1}\\ \boldsymbol{\theta}_{t|t-1}\end{pmatrix} + \mathbf{L}_t\left(\mathbf{x}_t - f_1\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right),\\
\mathbf{p}_{t+1|t} &= f_2\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right),\\
\boldsymbol{\theta}_{t+1|t} &= \boldsymbol{\theta}_{t|t},\\
\mathbf{L}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t\left(\mathbf{H}_t^T\mathbf{P}_{t|t-1}\mathbf{H}_t + \mathbf{R}\right)^{-1},\\
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{L}_t\mathbf{H}_t^T\mathbf{P}_{t|t-1},\\
\mathbf{P}_{t+1|t} &= \mathbf{D}_t\mathbf{P}_{t|t}\mathbf{D}_t^T + \mathbf{Q},
\end{aligned} \tag{23}$$

where

$$\begin{aligned}
\mathbf{p}_{t|t} &\triangleq \widetilde{E}\left[\mathbf{p}_t \mid \mathbf{x}_t, \mathbf{x}_{t-1}, \ldots\right],\\
\mathbf{p}_{t|t-1} &\triangleq \widetilde{E}\left[\mathbf{p}_t \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right],\\
\boldsymbol{\theta}_{t|t} &\triangleq \widetilde{E}\left[\boldsymbol{\theta}_t \mid \mathbf{x}_t, \mathbf{x}_{t-1}, \ldots\right],\\
\boldsymbol{\theta}_{t|t-1} &\triangleq \widetilde{E}\left[\boldsymbol{\theta}_t \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right]
\end{aligned} \tag{24}$$

are EKF terms that approximate the optimal "linear" MSE estimated values in the linearized case and $\mathbf{H}_t$ and $\mathbf{D}_t$ are the gradients for the first-order Taylor expansion needed to linearize the nonlinear state equations in (21)

$$\mathbf{H}_t = \begin{pmatrix}\left(\nabla_{\mathbf{p}_t} f_1\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right)^T\\ \left(\nabla_{\boldsymbol{\theta}_t} f_1\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right)^T\end{pmatrix}, \tag{25}$$

$$\mathbf{D}_t = \begin{pmatrix}\nabla_{\mathbf{p}_t} f_2\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right) & \nabla_{\boldsymbol{\theta}_t} f_2\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right)\\ \mathbf{0} & \mathbf{I}\end{pmatrix}, \tag{26}$$

respectively. Here, $\mathbf{L}_t$ is the gain of the EKF and $\mathbf{P}_t$ is the error variance of the augmented state. The complete set of equations in (23) defines the EKF update on the parameter vectors. We next consider the case when there is feedback.

3.2. Estimating the Unknown Parameters of the System with Feedback. For estimating the parameters of the feedback loop, that is, $\mathbf{c}_t$ (please see Figure 3), we have two different scenarios. In the first case, where we can control $\mathbf{w}$, we set $\mathbf{w} = \mathbf{0}$, estimate $\{\mathbf{F}, \mathbf{G}\}$, and then subsequently estimate $\mathbf{c}$ for fixed $\mathbf{w}$. For scenarios where the feedback loop is already present (or we cannot control it), that is, $\mathbf{w} \neq \mathbf{0}$, we need to estimate all the system parameters under the feedback loop. Naturally, in this case the estimation process is more prone to errors due to compounding effects of the feedback loop on the noise processes. We consider both cases separately.

Using (10) in (12), we get

$$\begin{aligned}
\mathbf{p}_{t+1} &= \mathbf{G}_t\mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{x}_t\\
&= \mathbf{G}_t\mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{F}_t\mathbf{p}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{n}_t.
\end{aligned} \tag{27}$$

Hence, the complete state-space description with causal loop is given by

$$\mathbf{p}_{t+1} = \left(\mathbf{G}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{F}_t\right)\mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{n}_t, \tag{28}$$

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{p}_t + \mathbf{n}_t. \tag{29}$$

In (29), $\mathbf{w}_t$ is known; however, all the parameters including $\mathbf{c}$ are unknown. We have two cases.

Case 1. Since we can control $\mathbf{w}$, we set $\mathbf{w} = \mathbf{0}$ and estimate $\boldsymbol{\theta}$ as $\widetilde{\mathbf{F}}_t$ and $\widetilde{\mathbf{G}}_t$ as in the case without feedback. Then, use these estimated parameters in (29) yielding

$$\begin{aligned}
\mathbf{p}_{t+1} &= \left(\widetilde{\mathbf{G}}_t + \mathbf{c}_t\mathbf{w}_t^T\widetilde{\mathbf{F}}_t\right)\mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t\mathbf{w}_t^T\mathbf{n}_t,\\
\mathbf{x}_t &= \widetilde{\mathbf{F}}_t\mathbf{p}_t + \mathbf{n}_t.
\end{aligned} \tag{30}$$

To estimate $\mathbf{c}_t$, we introduce an EKF framework by considering $\mathbf{c}_t$ as another state vector:

$$\mathbf{c}_{t+1} = \mathbf{c}_t + \boldsymbol{\rho}_t, \tag{31}$$

where $\boldsymbol{\rho}_t$ is the noise in the estimation process, yielding

$$\begin{aligned}
\mathbf{x}_t &= \widetilde{\mathbf{F}}_t\mathbf{p}_t + \mathbf{n}_t,\\
\begin{pmatrix}\mathbf{p}_{t+1}\\ \mathbf{c}_{t+1}\end{pmatrix} &= \begin{pmatrix}f_3\left(\mathbf{c}_t, \mathbf{p}_t\right)\\ \mathbf{c}_t\end{pmatrix} + \begin{pmatrix}\mathbf{v}_t\\ \boldsymbol{\rho}_t\end{pmatrix} + \begin{pmatrix}\mathbf{c}_t\mathbf{w}_t^T\\ \mathbf{0}\end{pmatrix}\mathbf{n}_t,
\end{aligned} \tag{32}$$

where

$$f_3\left(\mathbf{c}_t, \mathbf{p}_t\right) \triangleq \left(\widetilde{\mathbf{G}}_t + \mathbf{c}_t\mathbf{w}^T\widetilde{\mathbf{F}}_t\right)\mathbf{p}_t \tag{33}$$

is the corresponding nonlinearity in the system.

In the state update equation (32), unlike the previous EKF formulation, the process noise depends on $\mathbf{c}_t$ as $\mathbf{c}_t\mathbf{w}_t^T\mathbf{n}_t$, which is unknown and part of the estimated state vector. Hence, the EKF formulation is more involved.

After several steps, we derive the EKF equations to estimate the augmented states for this case as

$$\begin{aligned}
\begin{pmatrix}\mathbf{p}_{t|t}\\ \mathbf{c}_{t|t}\end{pmatrix} &= \begin{pmatrix}\mathbf{p}_{t|t-1}\\ \mathbf{c}_{t|t-1}\end{pmatrix} + \mathbf{L}_t\left(\mathbf{x}_t - \widetilde{\mathbf{F}}_t\mathbf{p}_{t|t-1}\right),\\
\mathbf{p}_{t+1|t} &= f_3\left(\mathbf{c}_{t|t}, \mathbf{p}_{t|t}\right) + \mathbf{S}_t\boldsymbol{\Omega}_t^{-1}\left(\mathbf{x}_t - \widetilde{\mathbf{F}}_t\mathbf{p}_{t|t-1}\right),\\
\mathbf{c}_{t+1|t} &= \mathbf{c}_{t|t},\\
\mathbf{L}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t\boldsymbol{\Omega}_t^{-1},\\
\mathbf{S}_t &= \mathbf{c}_{t|t-1}\mathbf{w}_t^T\mathbf{R},\\
\mathbf{P}_{t|t} &= \mathbf{H}_t^T\mathbf{P}_{t|t-1}\mathbf{H}_t + \mathbf{R},\\
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{L}_t\mathbf{H}_t^T\mathbf{P}_{t|t-1},\\
\mathbf{P}_{t+1|t} &= \mathbf{D}_t\mathbf{P}_{t|t-1}\mathbf{D}_t^T - \mathbf{B}_t\boldsymbol{\Omega}_t^{-1}\mathbf{B}_t^T + \widehat{\mathbf{Q}}_t,\\
\mathbf{B}_t &= \mathbf{D}_t\mathbf{P}_{t|t-1}\mathbf{H}_t + \begin{pmatrix}\mathbf{S}_t\\ \mathbf{0}\end{pmatrix},
\end{aligned} \tag{34}$$

where

$$\begin{aligned}\mathbf{p}_{t|t} &\triangleq \widetilde{E}\left[\mathbf{p}_t \mid \mathbf{x}_t, \mathbf{x}_{t-1}, \ldots\right],\\ \mathbf{p}_{t|t-1} &\triangleq \widetilde{E}\left[\mathbf{p}_t \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right],\\ \mathbf{c}_{t|t} &\triangleq \widetilde{E}\left[\mathbf{c}_t \mid \mathbf{x}_t, \mathbf{x}_{t-1}, \ldots\right],\\ \mathbf{c}_{t|t-1} &\triangleq \widetilde{E}\left[\mathbf{c}_t \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right]\end{aligned} \tag{35}$$

are EKF terms that approximate the optimal "linear" MSE estimated values in the linearized case and $\mathbf{H}_t$ and $\mathbf{D}_t$ are the gradients for the first-order Taylor expansion needed to linearize the nonlinear state equations in (32):

$$\begin{aligned}\mathbf{H}_t &= \begin{pmatrix}\left(\nabla_{\mathbf{p}_t}\left(\widetilde{\mathbf{F}}_t\mathbf{p}_t\right)\right)^T\\ \left(\nabla_{\mathbf{c}_t}\left(\widetilde{\mathbf{F}}_t\mathbf{p}_t\right)\right)^T\end{pmatrix} = \begin{pmatrix}\widetilde{\mathbf{F}}^T\\ \mathbf{0}\end{pmatrix},\\ \mathbf{D}_t &= \begin{pmatrix}\nabla_{\mathbf{p}_t} f_3\left(\mathbf{c}_{t|t}, \mathbf{p}_{t|t}\right) & \nabla_{\mathbf{c}_t} f_3\left(\mathbf{c}_{t|t}, \mathbf{p}_{t|t}\right)\\ \mathbf{0} & \mathbf{I}\end{pmatrix},\end{aligned} \tag{36}$$

respectively. Here, $\mathbf{L}_t$ is the gain of the EKF and $\mathbf{P}_t$ is the error variance of the augmented state.

To obtain an expression for $\widehat{\mathbf{Q}}_t$ in terms of $\mathbf{w}_t$, we define the composite error vector $\mathbf{b}_t$ for the state update equation so that

$$\widehat{\mathbf{Q}}_t = E\left[\mathbf{b}_t\mathbf{b}_t^T \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right] \tag{37}$$

with

$$\mathbf{b}_t \triangleq \begin{pmatrix}\mathbf{v}_t\\ \boldsymbol{\rho}_t\end{pmatrix} + \begin{pmatrix}\mathbf{c}_t\mathbf{w}_t^T\\ \mathbf{0}\end{pmatrix}\mathbf{n}_t. \tag{38}$$

After straightforward algebra, we get

$$\widehat{\mathbf{Q}}_t = \begin{pmatrix}\mathbf{Q} + \mathbf{w}_t^T\mathbf{R}\mathbf{w}_t\boldsymbol{\Gamma}_t & \mathbf{0}\\ \mathbf{0} & \mathbf{U}\end{pmatrix}, \tag{39}$$

where

$$\begin{aligned}\mathbf{U} &= E\left[\boldsymbol{\rho}_t\boldsymbol{\rho}_t^T\right],\\ \boldsymbol{\Gamma}_t &\triangleq \begin{pmatrix}\mathbf{0} & \mathbf{I}\end{pmatrix}\mathbf{p}_{t|t-1}\begin{pmatrix}\mathbf{0}\\ \mathbf{I}\end{pmatrix} + \mathbf{c}_{t|t-1}\mathbf{c}_{t|t-1}^T.\end{aligned} \tag{40}$$

These updates provide the complete EKF formulation with feedback. In the sequel, we introduce the complete estimation framework where we estimate all the parameters jointly.

Case 2. We can define a superset of parameters

$$\boldsymbol{\theta}_t \triangleq \left[\mathbf{G}_t(:); \mathbf{F}_t(:); \mathbf{c}_t\right] \tag{41}$$

and formulate an EKF framework for this augmented parameter vector with

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t, \tag{42}$$

which yields

$$\begin{aligned}\mathbf{x}_t &= f_4\left(\boldsymbol{\theta}_t, \mathbf{p}_t\right) + \mathbf{n}_t,\\ \begin{pmatrix}\mathbf{p}_{t+1}\\ \boldsymbol{\theta}_{t+1}\end{pmatrix} &= \begin{pmatrix}f_5\left(\boldsymbol{\theta}_t, \mathbf{p}_t\right)\\ \boldsymbol{\theta}_t\end{pmatrix} + \begin{pmatrix}\mathbf{v}_t\\ \boldsymbol{\varepsilon}_t\end{pmatrix} + \begin{pmatrix}\mathbf{c}_t\mathbf{w}_t^T\\ \mathbf{0}\end{pmatrix}\mathbf{n}_t,\end{aligned} \tag{43}$$

where

$$\begin{aligned}f_4\left(\boldsymbol{\theta}_t, \mathbf{p}_t\right) &\triangleq \mathbf{F}_t\mathbf{p}_t,\\ f_5\left(\boldsymbol{\theta}_t, \mathbf{p}_t\right) &\triangleq \left(\mathbf{G}_t - \mathbf{c}_t\mathbf{w}^T\mathbf{F}_t\right)\end{aligned} \tag{44}$$

are the corresponding nonlinear equations so that we require EKF.

After some algebra, we get the complete EKF equations as

$$\begin{aligned}\begin{pmatrix}\mathbf{p}_{t|t}\\ \boldsymbol{\theta}_{t|t}\end{pmatrix} &= \begin{pmatrix}\mathbf{p}_{t|t-1}\\ \boldsymbol{\theta}_{t|t-1}\end{pmatrix} + \mathbf{L}_t\left(\mathbf{x}_t - f_4\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right),\\ \mathbf{p}_{t+1|t} &= f_5\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right)\\ &\quad + \mathbf{S}_t\boldsymbol{\Omega}_t^{-1}\left(\mathbf{x}_t - f_4\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right),\\ \boldsymbol{\theta}_{t+1|t} &= \boldsymbol{\theta}_{t|t},\\ \mathbf{L}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t\boldsymbol{\Omega}_t^{-1},\\ \mathbf{S}_t &= \begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{I}\end{pmatrix}\boldsymbol{\theta}_{t|t-1}\mathbf{w}_t^T\mathbf{R},\\ \boldsymbol{\Omega}_t &= \mathbf{H}_t^T\mathbf{P}_{t|t-1}\mathbf{H}_t + \mathbf{R},\\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{L}_t\mathbf{H}_t^T\mathbf{P}_{t|t-1},\\ \mathbf{P}_{t+1|t} &= \mathbf{D}_t\mathbf{P}_{t|t-1}\mathbf{D}_t^T - \mathbf{B}_t\boldsymbol{\Omega}_t^{-1}\mathbf{B}_t^T + \widehat{\mathbf{Q}}_t,\\ \mathbf{B}_t &= \mathbf{D}_t\mathbf{P}_{t|t-1}\mathbf{H}_t + \begin{pmatrix}\mathbf{S}_t\\ \mathbf{0}\end{pmatrix},\end{aligned} \tag{45}$$

where

$$\begin{aligned}\mathbf{H}_t &= \begin{pmatrix}\left(\nabla_{\mathbf{p}_t} f_4\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right)^T\\ \left(\nabla_{\boldsymbol{\theta}_t} f_4\left(\boldsymbol{\theta}_{t|t-1}, \mathbf{p}_{t|t-1}\right)\right)^T\end{pmatrix},\\ \mathbf{D}_t &= \begin{pmatrix}\nabla_{\mathbf{p}_t} f_5\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right) & \nabla_{\boldsymbol{\theta}_t} f_5\left(\boldsymbol{\theta}_{t|t}, \mathbf{p}_{t|t}\right)\\ \mathbf{0} & \mathbf{I}\end{pmatrix}.\end{aligned} \tag{46}$$

To obtain an expression for $\widehat{\mathbf{Q}}_t$ in terms of $\mathbf{w}_t$, we define the composite error vector $\mathbf{b}_t$ for the state update equation so that

$$\widehat{\mathbf{Q}}_t = E\left[\mathbf{b}_t\mathbf{b}_t^T \mid \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\right] \tag{47}$$

with

$$\mathbf{b}_t \triangleq \begin{pmatrix}\mathbf{v}_t\\ \boldsymbol{\varepsilon}_t\end{pmatrix} + \begin{pmatrix}\mathbf{c}_t\mathbf{w}_t^T\\ \mathbf{0}\end{pmatrix}\mathbf{n}_t. \tag{48}$$

After straightforward algebra, we get

$$\widehat{\mathbf{Q}}_t = \begin{pmatrix} \mathbf{Q} + \mathbf{w}_t^T \mathbf{R} \mathbf{w}_t \boldsymbol{\Gamma}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_t \end{pmatrix}, \tag{49}$$

where

$$\begin{aligned} \mathbf{Q} &= E\left[\boldsymbol{\nu}_t \boldsymbol{\nu}_t^T\right], \\ \mathbf{U} &= E\left[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^T\right], \\ \mathbf{R} &= E\left[\mathbf{n}_t \mathbf{n}_t^T\right], \\ \boldsymbol{\Gamma}_t &\triangleq \begin{pmatrix} \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{p}_{t|t-1} \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \\ &\quad + \begin{pmatrix} \mathbf{0} & \mathbf{I} \end{pmatrix} \boldsymbol{\theta}_{t|t-1} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \boldsymbol{\theta}_{t|t-1}^T \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix}. \end{aligned} \tag{50}$$

Given that the system parameters are estimated through the EKF formulation, we next introduce learning algorithms on $\mathbf{w}_t$ in order to change the behavior of the users in a desired manner.

4. Designing a Causal Inference System to Tune User Preferences

After the parameters are estimated through methods described in the previous sections, the complete system framework is given by

$$\begin{aligned} \mathbf{x}_t &= \mathbf{F}_t \mathbf{p}_t + \mathbf{n}_t, \\ \mathbf{p}_{t+1} &= \left(\mathbf{G}_t + \mathbf{c}_t \mathbf{w}_t^T \mathbf{F}_t\right) \mathbf{p}_t + \mathbf{v}_t + \mathbf{c}_t \mathbf{w}_t^T \mathbf{n}_t, \end{aligned} \tag{51}$$

with the estimated

$$\left\{\mathbf{F}_t = \widetilde{\mathbf{F}}_t, \mathbf{G}_t = \widetilde{\mathbf{G}}_t, \widetilde{\mathbf{c}}_t = \mathbf{c}_t\right\}. \tag{52}$$

Our goal in this section is to design $\mathbf{w}_t$ such that the sequence of preferences $\mathbf{p}_t$ are tuned towards a desired sequence of preferences $\mathbf{q}_t$; for example, one can desire to sway the preferences of a user to a certain product.

In order to tune the user preferences, we design $\mathbf{w}_t$ so that the difference between the preferences $\mathbf{p}_t$ and the desired $\mathbf{q}_t$ is minimized. We define this difference as the loss between the preferences and desired vectors as

$$\sum_{k=1}^{t} l\left(\mathbf{p}_k, \mathbf{q}_k\right), \tag{53}$$

where $l(\cdot)$ is any differentiable loss function. As an example, for the square error loss, this yields

$$\sum_{k=1}^{t} \left\|\mathbf{p}_k - \mathbf{q}_k\right\|^2. \tag{54}$$

To minimize the difference between these two sequences, we introduce a stochastic gradient approach where $\mathbf{w}_t$ is learned in a sequential manner. In the stochastic gradient approach, we have

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \mu \nabla_{\mathbf{w}} l\left(\mathbf{p}_k, \mathbf{q}_k\right), \tag{55}$$

where $\mu > 0$ is an appropriate learning rate coefficient. The learning rate coefficient is usually selected as time varying with two conditions:

$$\begin{aligned} &\mu_t \longrightarrow 0 \quad \text{as } t \longrightarrow \infty, \\ &\sum_{k=1}^{t} \mu_k \longrightarrow \infty \quad \text{as } t \longrightarrow \infty; \end{aligned} \tag{56}$$

for example, $\mu_t = 1/t$.

If these two conditions are met, then the estimated parameters $\mathbf{w}_t$ through the gradient approach will converge to the optimal $\mathbf{w}$ (provided that such an optimal point exists) [21]. To facilitate the analysis, we set

$$l\left(\mathbf{p}_k, \mathbf{q}_k\right) = \left\|\mathbf{p}_k - \mathbf{q}_k\right\|^2 \tag{57}$$

and get

$$\begin{aligned} \mathbf{w}_{t+1} &= \mathbf{w}_t - \mu \nabla_{\mathbf{w}_t} \left\|\mathbf{p}_k - \mathbf{q}_k\right\|^2 \\ &= \mathbf{w}_t - 2\mu \left(\nabla_{\mathbf{w}_t} \mathbf{p}_t\right) \left(\mathbf{p}_t - \mathbf{q}_t\right). \end{aligned} \tag{58}$$

In (58), since $\mathbf{p}_t$ is unknown, we use $\mathbf{p}_{t|t-1}$ from the causal loop case, that is, with feedback, and get

$$\mathbf{w}_{t+1} = \mathbf{w}_t - 2\mu \left(\nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1}\right) \left(\mathbf{p}_{t|t-1} - \mathbf{q}_t\right). \tag{59}$$

To get

$$\nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1}, \tag{60}$$

we use the EKF recursion as

$$\begin{aligned} \mathbf{p}_{t|t-1} &= \left(\mathbf{G}_t + \mathbf{c}_t \mathbf{w}_t^T \mathbf{F}_t\right) \\ &\quad \cdot \left(\mathbf{p}_{t-1|t-2} + \mathbf{L}_{t-1}\left[\mathbf{x}_{t-1} - \mathbf{H}_{t-1} \mathbf{p}_{t-1|t-2}\right]\right), \\ \mathbf{p}_{t|t-1} &= \mathbf{K}_t \mathbf{p}_{t|t-1} + \mathbf{M}_t, \end{aligned} \tag{61}$$

where

$$\begin{aligned} \mathbf{K}_t &= \left(\mathbf{G}_t + \mathbf{c}_t \mathbf{w}_t^T \mathbf{F}_t\right)\left(\mathbf{I} - \mathbf{L}_{t-1} \mathbf{H}_{t-1}\right), \\ \mathbf{M}_t &= \left(\mathbf{G}_t + \mathbf{c}_t \mathbf{w}_t^T \mathbf{F}_t\right) \mathbf{L}_t x_{t-1}. \end{aligned} \tag{62}$$

Using (61), we get a recursive update on the gradient as

$$\nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1} = \nabla_{\mathbf{w}_t} \mathbf{K}_t \mathbf{p}_{t|t-1} + \mathbf{K}_t \nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1} \nabla_{\mathbf{w}_t} \mathbf{M}_t, \tag{63}$$

From (59), (61), and (63), we get the complete recursive update as

$$\boxed{\begin{aligned} \mathbf{w}_{t+1} &= \mathbf{w}_t - 2\mu \left(\nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1}\right)\left(\mathbf{p}_{t|t-1} - \mathbf{q}_t\right) \\ \mathbf{p}_{t|t-1} &= \mathbf{K}_t \mathbf{p}_{t|t-1} + \mathbf{M}_t \\ \nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1} &= \nabla_{\mathbf{w}_t} \mathbf{K}_t \mathbf{p}_{t|t-1} + \mathbf{K}_t \nabla_{\mathbf{w}_t} \mathbf{p}_{t|t-1} \nabla_{\mathbf{w}_t} \mathbf{M}_t \end{aligned}}. \tag{64}$$

This completes the derivation of the stochastic gradient update for online learning of the tuning regression vector.

5. Experiments

In this section, we share our simulation results to show that estimated parameters of the system converge to the real values, proving that a system can be designed with the right parameters which allows a sequence of actions or interventions to tune the preferences of a user in a desired manner. Since our goal is mainly to establish a pathway to the possibility of designing a system that can steer user preferences in a desired manner, we consider our basic simulation set to be sufficient based on the mathematical proof we provided in the form of EKF formulations. The true parameters of the system are known to us since we are running our experiments in the form of simulations. Specifically, the preferences of the user, which are not directly observable in real life, are known in case of simulations. We run simulations for the EKF formulations we derived in the previous sections to show that our estimation of the preferences converges to the real preference values. We illustrate the convergence of our algorithms under different scenarios.

In the first scenario, we have the case where the corresponding system has no feedback. As the true system, we choose a second-order linear state-space model, where $\mathbf{G} = 0.95\mathbf{I}$ and $\mathbf{F} = \mathbf{I}$ with $\mathbf{Q} = 3 \times 10^{-3}\mathbf{I}$ and $\mathbf{R} = 3 \times 10^{-3}\mathbf{I}$. For the EKF formulation, we choose two different variances for ε_t, for example, 10^{-3} and 10^{-4}, to demonstrate the effect of this design parameter on the system. We emphasize that neither $\mathbf{F}$ or $\mathbf{G}$ are known; hence, as long as the system is observable, particular choices of $\mathbf{F}$ and $\mathbf{G}$ only change the convergence speed and the final MSE. However, we choose $\mathbf{F}$ to make the system stable.

In Figure 4, we plot the square error difference between the estimated preferences and the real preferences

$$\operatorname{tr} E\left[\left\|\mathbf{p}_t - \mathbf{p}_{t|t-1}\right\|^2\right] \tag{65}$$

with respect to the number of iterations, where we produce the MSE curves after averaging over 100 independent trials. We also plot the cumulative MSE normalized with respect to time, that is,

$$\frac{\sum_{k=1}^{t} \operatorname{tr} E\left[\left\|\mathbf{p}_t - \mathbf{p}_{t|t-1}\right\|^2\right]}{t}, \tag{66}$$

to show that as the iteration count increases, the averaged MSE steadily converges. The plot includes both the average MSE and the cumulative MSE normalized in time for estimation of $\mathbf{F}$ and $\mathbf{G}$. We observe that the estimation of $\mathbf{F}$ and $\mathbf{G}$ is more prone to errors due to the multiplicative uncertainty, single observation, and state update equations. However, both the estimated preferences vectors as well as the system parameters converge.

In the second set of experiments, we have feedback present; that is, $\mathbf{w} \neq \mathbf{0}$. For this case, we now have similar parameters as in the first set of experiments, except $\mathbf{G} = 0.9\mathbf{I}$ to give more decay due to presence of feedback. For this case, we choose two different scenarios, where $\mathbf{w}_t$ and $\mathbf{c}_t$ are fixed or randomly chosen provided that the overall system stays stable after the feedback; that is, $(\mathbf{G}+\mathbf{c}\,\mathbf{w}^T\,\mathbf{F})$ corresponds to a stable system. Note that this can be always forced by choosing an appropriate $\mathbf{w}$. However, we choose randomly initialized $\mathbf{w}$ to avoid any bias in our experiments. Here, although $\mathbf{w}$ is known to us, the feedback amount $\mathbf{c}$ and the hidden preferences are unknown. In Figure 5, we plot the MSE between the

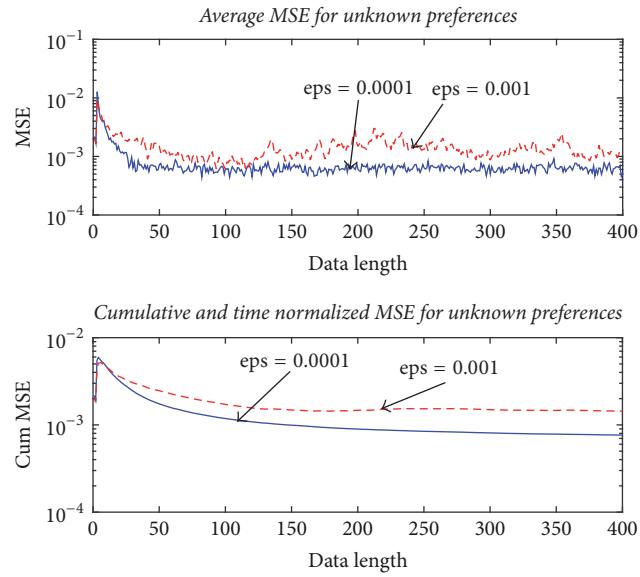


Figure 4: Estimation of the underlying preferences vector when there is no feedback. The results are averaged over 100 independent trials. Here, we have no feedback and parameters of both the state equation and the observation equation are unknown. The results are shown for two different noise variances for the EKF formulation.

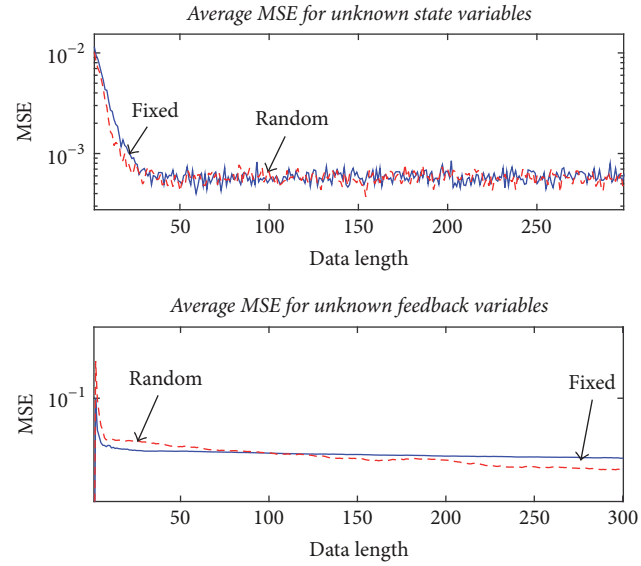


Figure 5: Estimation of the underlying vector of preferences and the feedback parameters when there is feedback. The results are averaged over 100 independent trials. Two different configurations are simulated for the feedback as well as for the linear control parameters, for example, the fixed and random initial cases. For both scenarios, our estimation process converges to the true underlying processes.

estimated preference vectors and the true ones. We observe from these simulations that although the feedback produces a multiplicative uncertainty in the state equation and greatly enhances the nonlinearity in the update equation, we are able to recover the true values through the EKF formulation. We observe that although due to feedback we have more colored noise in the state equation, we recover true values due to the whitening effects of the EKF. The MSE errors between the estimated feedback and the true one are plotted, where the MSE curves are produced after 100 independent realizations.

6. Conclusions

In this paper, we model the effects of the machine learning algorithms such as recommendation engines on users through a causal feedback loop. To this end, we introduce a complete state-space formulation modeling: (1) evolution of preference vectors, (2) observations generated by users, and (3) the causal feedback effects of the actions of machine learning algorithms on the system. All these parameters are jointly optimized through an Extended Kalman Filtering framework. We introduce algorithms to estimate the unknown system parameters with and without feedback. In both cases, all the parameters are estimated jointly. We emphasize that we provide a complete set of equations covering all the possible scenarios. To tune the preferences of users towards a desired sequence, we also introduce a linear feedback and introduce an optimization framework using stochastic gradient descent algorithm. Unlike previous works that only use the observations to predict certain desired quantities, we specifically design outputs to "update" the internal state of the system in a desired manner. Through a set of experiments, we demonstrate the convergence behavior of our proposed algorithms in different scenarios.

We consider our work as a significant theoretical first step in designing a system with the right parameters which allows a sequence of actions or interventions to tune the preferences of a user in a desired manner. We emphasize that the main goal of our study is to establish a pathway to designing such a system. We achieve this by first providing mathematical proof and then through a basic set of simulations.

A next step in future studies can be to make the system more stable and also to make the design process easy and practical for system designers. Further analysis on the convergence of the system and more simulations, experiments, and numerical analyses are needed to take our results to the next level. A direct comparison to previous studies is not possible for this first step of our study since, to the best of our knowledge, this is the first time a task of this nature is being undertaken. Our main success criterion is the fact that estimated parameters converge to the real parameter values. However, as our framework evolves, we will be able to track its relative performance.

Another area of focus for future studies is the optimal selection of action sequences. This can be particularly challenging since user preferences can change over time due to the abundance of new products and services. Algorithms to optimally select actions may require online learning and decision making in real time to accommodate these changes.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors would like to thank Koc University Graduate School of Social Sciences and Humanities for their support. This work was also supported by the BAGEP Award of the Science Academy.

References

[1] V. Gupta, D. Varshney, H. Jhamtani, D. Kedia, and S. Karwa, "Identifying purchase intent from social posts," in *Proceedings of the 8th International Conference on Weblogs and Social Media (ICWSM '14)*, June 2014.

[2] D. Ruta, "Automated trading with machine learning on big data," in *Proceedings of the 3rd IEEE International Congress on Big Data (BigData Congress '14)*, pp. 824–830, July 2014.

[3] Y. Wang and P. M. Djurić, "Social learning with heterogeneous agents and sequential decision making," *Digital Signal Processing*, vol. 47, pp. 17–24, 2015.

[4] L. Bottou and Y. Le Cun, "On-line learning for very large data sets," *Applied Stochastic Models in Business and Industry*, vol. 21, no. 2, pp. 137–151, 2005.

[5] L. Bottou and O. Bousquet, "The tradeoffs of large scale learning," *In Advances in Neural Information Processing (NISP)*, pp. 1–8, 2007.

[6] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, "How much can behavioral targeting help online advertising?" in *Proceedings of the 18th International World Wide Web Conference (WWW '09)*, pp. 261–270, April 2009.

[7] H. Peng, D. Liang, and C. Choi, "Evaluating parallel logistic regression models," in *Proceedings of the 2013 IEEE International Conference on Big Data (Big Data '13)*, pp. 119–126, October 2013.

[8] L. Bottou, J. Peters, J. Quiñonero-Candela et al., "Counterfactual reasoning and learning systems: the example of computational advertising," *Journal of Machine Learning Research*, vol. 14, pp. 3207–3260, 2013.

[9] Y. C. Sübakan, B. Kurt, A. T. Cemgil, and B. Sankur, "Probabilistic sequence clustering with spectral learning," *Digital Signal Processing: a Review Journal*, vol. 29, no. 1, pp. 1–19, 2014.

[10] Y. Achbany, I. J. Jureta, S. Faulkner, and F. Fouss, "Continually learning optimal allocations of services to tasks," *IEEE Transactions on Services Computing*, vol. 1, no. 3, pp. 141–154, 2008.

[11] M. Jahrer, A. Töscher, and R. Legenstein, "Combining predictions for accurate recommender systems," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 693–701, July 2010.

[12] A. Töscher, M. Jahrer, and R. Legenstein, "Improved neighborhood-based algorithms for large-scale recommender systems," in *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition (NETFLIX '08)*, August 2008.

[13] D. Chan, R. Ge, O. Gershony, T. Hesterberg, and D. Lambert, "Evaluating online ad campaigns in a pipeline: causal models at scale," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 7–15, July 2010.

[14] R. Epstein and R. E. Robertson, "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 33, pp. E4512–E4521, 2015.

[15] A. A. Salah, B. Lepri, F. Pianesi, and A. S. Pentland, "Human behavior understanding for inducing behavioral change: application perspectives," in *Proceedings of the 2nd International Workshop on Human Behavior Understanding*, vol. 7065, pp. 1–15, 2011.

[16] T. Z. Zarsky, *Thinking outside The Box: considering Transparency, Anonymity, and Pseudonymity as Overall Solutions to the Problems of Information Privacy in the Internet Society*, vol. 58, University of Miami Law Review, 2004.

[17] P. Wang, D. Yin, M. Meytlis, J. Yang, and Y. Chang, "Rethink targeting: detect 'smart cheating' in online advertising through causal inference," in *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*, pp. 133-134, May 2015.

[18] P. Wang, W. Sun, D. Yin, J. Yang, and Y. Chang, "Robust tree-based causal inference for complex ad effectiveness analysis," in *Proceedings of the 8th ACM International Conference on Web Search and Data Mining (WSDM '15)*, pp. 67–76, February 2015.

[19] K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, and S. L. Scott, "Inferring causal impact using bayesian structural time-series models," *Annals of Applied Statistics*, vol. 9, no. 1, pp. 247–274, 2015.

[20] W. Sun, P. Wang, D. Yin, J. Yang, and Y. Chang, "Causal inference via sparse additive models with application to online advertising," in *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI '15)*, 2015.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, USA, 2006.

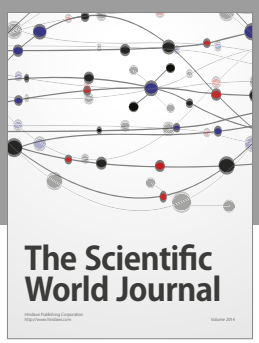
[22] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, New Jersey, NJ, USA, 1979.

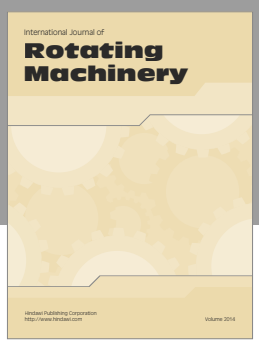
[23] S. S. Kozat, A. C. Singer, and G. C. Zeitler, "Universal piecewise linear prediction via context trees," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3730–3745, 2007.

[24] N. D. Vanli and S. S. Kozat, "A comprehensive approach to universal piecewise nonlinear regression based on trees," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5471–5486, 2014.

[25] A. C. Singer, S. S. Kozat, and M. Feder, "Universal linear least squares prediction: upper and lower bounds," *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2354–2362, 2002.

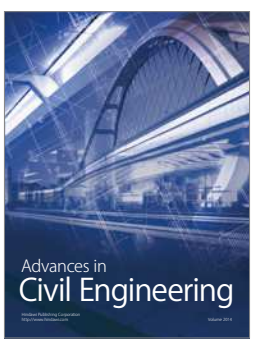


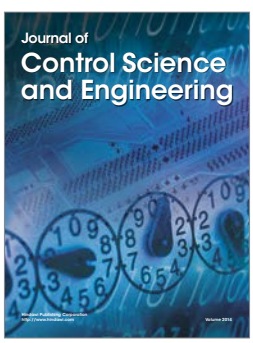














Hindawi
Submit your manuscripts at
https://www.hindawi.com