Incentive Schemes for Privacy-Sensitive Consumers

Chong Huang[1](✉), Lalitha Sankar[1], and Anand D. Sarwate[2]

[1] Arizona State University, Tempe, USA
{chong.huang,lalithasankar}@asu.edu
[2] Rutgers, The State University of New Jersey, New Brunswick, USA
anand.sarwate@rutgers.edu

Abstract. Businesses (*retailers*) often offer personalized advertisements (*coupons*) to individuals (*consumers*). While proving a customized shopping experience, such coupons can provoke strong reactions from consumers who feel their privacy has been violated. Existing models for privacy try to *quantify* privacy risk but do not capture the subjective experience and heterogeneous *expression* of privacy-sensitivity. We use a Markov decision process (MDP) model for this problem. Our model captures different consumer privacy sensitivities via a time-varying state, different coupon types via an action set for the retailer, and a cost for perceived privacy violations that depends on the action and state. The simplest version of our model has two states ("Normal" and "Alerted"), two coupons (targeted and untargeted), and consumer behavior dynamics known to the retailer. We show that the optimal coupon-offering strategy for a retailer that wishes to minimize its expected discounted cost is a stationary threshold-based policy. The threshold is a function of all model parameters: the retailer offers a targeted coupon if their belief that the consumer is in the "Alerted" state is below the threshold. We extend our model and results to consumers with multiple privacy-sensitivity states as well as coupon-dependent state transition probabilities.

Keywords: Privacy · Markov decision processes · Retailer-consumer interaction · Optimal policies

1 Introduction

Programs such as retailer "loyalty cards" allow companies to automatically track a customer's financial transactions, purchasing behavior, and preferences. They can then use this information to offer customized incentives, such as discounts on related goods. Consumers may benefit from retailer's knowledge by using more of these targeted discounts or coupons while shopping. However, the coupon offer may imply that the retailer has learned something sensitive or private about the consumer (for example, a pregnancy [1]) – such violations may make consumers skittish about purchasing from such retailers.

MHR Khouzani et al. (Eds.): GameSec 2015, LNCS 9406, pp. 358–369, 2015.
DOI: 10.1007/978-3-319-25594-1_21

However, modeling the privacy-sensitivity of a consumer is not always straightforward: widely-studied models for quantifying privacy risk using differential privacy [2] or information theory [3] do not capture the subjective experience and heterogeneous *expression* of consumer privacy. We introduce a framework to model the consumer-retailer interaction problem and better understand how retailers can develop coupon-offering policies that balances their revenue objectives while being sensitive to consumer privacy concerns. The main challenge for the retailer is that the consumer's responses to coupons are not known *a priori*; furthermore, consumers do not "add noise" to their purchasing behavior as a mechanism to stay private. Rather, the offer of a coupon may provoke a reaction from the consumer, ranging from "indifferent" through "partially concerned" to "creeped out." This reaction is mediated by the consumer's sensitivity level to privacy violations, and it is these levels that we seek to model via a Markov decision process. In particular, the sensitivity of the consumers are often revealed indirectly to the retailer through their purchasing patterns. We capture these aspects in our model and summarize our main contributions below.

Main Contributions: We propose a partially-observed Markov decision process (POMDP) model for this problem in which the consumer's state encodes their privacy sensitivity, and the retailer can offer different levels of privacy-violating coupons. The simplest instance of our model is one with two states for the consumer, denoted as "Normal" and "Alerted," and two types of coupons: untargeted *low privacy* (LP) or targeted *high privacy* (HP). At each time, the retailer may offer a coupon and the consumer transitions from one state to another according to a Markov chain that is independent of the offered coupon. The retailer suffers a cost that depends both on the type of coupon offered and the state of the consumer. The costs reflect the advantage of offering targeted HP coupons relative to untargeted LP ones while simultaneously capturing the risk of doing so when the consumer is already "Alerted".

Under the assumption that the retailer (via surveys or prior knowledge) knows the statistics of the consumer Markov process, i.e., the likelihoods of becoming "Alerted" and staying "Alerted", and a belief about the initial consumer state, we study the problem of determining the optimal coupon-offering policy that the retailer should adopt to minimize the long-term discounted costs of offering coupons. We show that the optimal stationary policy exists and it is a threshold on the probability of the consumer being alerted; this threshold is a function of all the model parameters. The simple model above is extended to multiple consumer states and coupon-dependent transitions. We model the latter via two Markov processes for the consumer, one for each type (HP or LP) of coupon such that a persnickety consumer who is easily "Alerted" will be more likely to do so when offered an HP (relative to LP) coupon. Our structural result (a stationary optimal policy) holds for multiple states and coupon-dependent transitions. While the MDP model used in this paper is simple, its application to the problem of privacy cost minimization with privacy-sensitive consumers is novel. In the conclusion we describe several other interesting avenues for future work. Our results use many fundamental tools and techniques from the theory of

MDPs through appropriate and meaningful problem modeling. We briefly review the related literature in consumer privacy studies as well as MDPs.

Related Work: Several economic studies have examined consumer's attitudes towards privacy via surveys and data analysis including studies on the benefits and costs of using private data (e.g., Aquisti and Grossklags in [4]). On the other hand, formal methods such as differential privacy are finding use in modeling the value of private data for market design [5] and for the problem of partitioning goods with private valuation function amongst the agents [6]. In these models the goal is to elicit private information from individuals. Venkitasubramaniam [7] recently used an MDP model to study data sharing in control systems with time-varying state. He explicitly quantifies privacy risk in terms of equivocation, an information-theoretic measure, and his objective is to minimize the weighted sum of the utility (benefit) that the system achieves by sharing data (e.g., with a data collector) and the privacy risk. In our work we do not quantify *privacy risk* directly; instead the retailer learns about the *privacy-sensitivity* of the consumer indirectly through the cost feedback. Our MDP's state space is the privacy sensitivity of the consumer. To the best of our knowledge, models capturing this aspect of consumer-retailer interactions and the related privacy issues have not been studied before; in particular, our work focuses on explicitly considering the consequence to the retailer of the consumers' awareness of privacy violations.

Markov decision processes (MDPs) have been widely used for decades across many fields [8]; in particular, our formal model is related to problems in control with communication constraints [9,10] where state estimation has a cost. However, our costs are action and state dependent and we consider a different optimization problem. Classical state-search problems [11,12] also have optimal threshold policies; however the retailer's objective in our model is to minimize cost, and not necessarily estimate the consumer state. Our model is most similar to Ross's model of product quality control with deterioration [13], which was more recently used by Laourine and Tong to study the Gilbert-Elliot channel in wireless communications [14], in which the channel has two states and the transmitter has two actions (to transmit or not). We cannot apply their results directly due to our different cost structure, but use ideas from their proofs. Furthermore, we go beyond these works to study privacy-utility tradeoffs in consumer-retailer interactions with more than two states and action-dependent transition probabilities. We apply more general MDP analysis tools to address our formal behavioral model for privacy-sensitive consumers.

2 System Model

We model interactions between a retailer and a consumer via a discrete-time system (see Fig. 1). At each time t, the consumer has a discrete-valued state and the retailer may offer one of two coupons: high privacy risk (HP) or low privacy risk (LP). The consumer responds by imposing a cost on the retailer that depends on the coupon offered and its own state. For example, a consumer who is "alerted" (privacy-aware) may respond to an HP coupon by refusing to shop at the retailer. The retailer's goal is to decide which type of coupon to offer at each time t to minimize its cost.

2.1 Consumer Model

Modeling Assumption 1 ***(Consumer's State).*** *We assume the consumer is in one of a finite set of states that determine their response to coupons – each state corresponds to a type of consumer behavior in terms of purchasing. The consumer's state evolves according to a Markov process.*

For this paper, we primarily focus on the two-state case; the consumer may be Normal or Alerted. Later we will extend this model to multiple consumer states. The consumer state at time t is denoted by $G_t \in \{\mathsf{Normal}, \mathsf{Alerted}\}$. If a consumer is in Normal state, the consumer is very likely to use coupons to make purchases. However, in the Alerted state, the consumer is less likely to use coupons, since it is more cautious about revealing information to the retailer. The evolution of the consumer state is modeled as an infinite-horizon discrete time Markov chain (Fig. 1). The consumer starts out in a random initial state unknown to the retailer and the transition of the consumer state is independent of the action of the retailer. A *belief state* is a probability distribution over possible states in which the consumer could be. The belief of the consumer being in Alerted state at time t is denoted by p_t. We define $\lambda_{N,A} = Pr[G_t = \mathsf{Alerted}|G_{t-1} = \mathsf{Normal}]$ to be the transition probability from Normal state to Alerted state and $\lambda_{A,A} = Pr[G_t = \mathsf{Alerted}|G_{t-1} = \mathsf{Alerted}]$ to be the probability of staying in Alerted state when the previous state is also Alerted. The transition matrix $\mathbf{\Lambda}$ of the Markov chain can be written as

$$\mathbf{\Lambda} = \begin{pmatrix} 1 - \lambda_{N,A} & \lambda_{N,A} \\ 1 - \lambda_{A,A} & \lambda_{A,A} \end{pmatrix}. \tag{1}$$

We assume the transition probabilities are known to the retailer; this may come from statistical analysis such as a survey of consumer attitudes. The one step transition function, defined by $T(p_t) = (1 - p_t)\lambda_{N,A} + p_t\lambda_{A,A}$, represents the belief that the consumer is in Alerted state at time $t + 1$ given p_t, the Alerted state belief at time t.

Modeling Assumption 2 ***(State Transitions).*** *Consumers have an inertia in that they tend to stay in the same state. Moreover, once consumers feel their privacy is violated, it will take some time for them to come back to* Normal *state.*

To guarantee Assumption 2 we consider transition matrices in (1) satisfying $\lambda_{A,A} \geq 1 - \lambda_{A,A}$, $1 - \lambda_{N,A} \geq \lambda_{N,A}$, and $\lambda_{N,A} \geq 1 - \lambda_{A,A}$. Thus, by combining the above three inequalities, we have $\lambda_{A,A} \geq \lambda_{N,A}$.

2.2 Retailer Model

At each time t, the retailer can take an *action* by offering a coupon to the consumer. We define the action at time t to be $u_t \in \{\mathsf{HP}, \mathsf{LP}\}$, where HP denotes

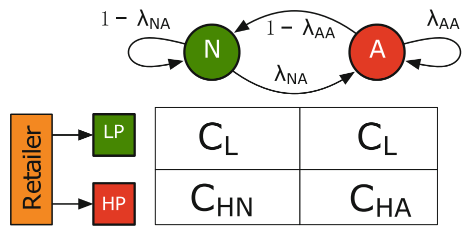


Fig. 1. Markov state transition model for a two-state consumer.

offering a high privacy risk coupon (e.g. a targeted coupon) and LP denotes offering a low privacy risk coupon (e.g. a generic coupon). The retailer's utility is modeled by a *cost* (negative revenue) which depends on both the consumer's state and the type of coupon being offered. If the retailer offers an LP coupon, it suffers a cost C_L independent of the consumer's state: offering LP coupons does not reveal anything about the state. However, if the retailer offers an HP coupon, then the cost is C_{HN} or C_{HA} depending on whether the consumer's state is Normal or Alerted. Offering an HP (high privacy risk, targeted) coupon to a Normal consumer should incur a low cost (high reward), but offering an HP coupon to an Alerted consumer should incur a high cost (low reward) since an Alerted consumer is privacy-sensitive. Thus, we assume $C_{HN} \leq C_L \leq C_{HA}$.

Under these conditions, the retailer's objective is to choose u_t at each time t to minimize the total cost incurred over the entire time horizon. The HP coupon reveals information about the state through the cost, but is risky if the consumer is alerted, creating a tension between cost minimization and acquiring state information.

2.3 The Minimum Cost Function

We define $C(p_t, u_t)$ to be the expected cost acquired from an individual consumer at time t where p_t is the probability that the consumer is in Alerted state and u_t is the retailer's action:

$$C(p_t, u_t) = \begin{cases} C_L & \text{if } u_t = \mathsf{LP} \\ (1-p_t)C_{HN} + p_t C_{HA} & \text{if } u_t = \mathsf{HP} \end{cases}. \tag{2}$$

Since the retailer knows the consumer state from the incurred cost only when an HP coupon is offered, the state of the consumer may not be directly observable to the retailer. Therefore, the problem is actually a Partially Observable Markov Decision Process (POMDP) [15].

We model the cost of violating a consumer's privacy as a short term effect. We adopt a discounted cost model with discount factor $\beta \in (0, 1)$. At each time t, the retailer has to choose which action u_t to take in order to minimize the expected discounted cost over infinite time horizon. A policy π for the retailer is a rule that selects a coupon to offer at each time. Given that the belief of the consumer being in Alerted state at time t is p_t and the policy is π, the infinite-horizon discounted cost starting from t is

$$V_\beta^{\pi,t}(p_t) = \mathbb{E}_\pi\left[\sum_{i=t}^{\infty}\beta^i C(p_i, A_i)|p_t\right], \tag{3}$$

where $\mathbb{E}_\pi$ indicates the expectation over the policy π. The objective of the retailer is equivalent to minimizing the discounted cost over all possible policies. We define the minimum cost function starting from time t over all policies to be

$$V_\beta^t(p_t) = \min_\pi V_\beta^{\pi,t}(p_t) \text{ for all } p_t \in [0,1]. \tag{4}$$

We define p_{t+1} to be the belief of the consumer being in Alerted state at time $t+1$. The minimum cost function $V_\beta^t(p_t)$ satisfies the Bellman equation [15]:

$$V_\beta^t(p_t) = \min_{u_t\in\{\mathsf{HP},\mathsf{LP}\}}\{V_{\beta,u_t}^t(p_t)\} \tag{5}$$

$$V_{\beta,u_t}^t(p_t) = \beta^t C(p_t, u_t) + V_\beta^{t+1}(p_{t+1}|p_t, u_t). \tag{6}$$

An optimal policy is *stationary* if it is a deterministic function of states, i.e., the optimal action at a particular state is the optimal action in this state at all times. We define $\mathcal{P} = \{[0,1]\}$ to be the belief space and $\mathcal{U} = \{\mathsf{LP}, \mathsf{HP}\}$ to be the action space. In the context of our model, the optimal stationary policy is a deterministic function mapping $\mathcal{P}$ into $\mathcal{U}$. Since the problem is an infinite-horizon, finite state, and finite action MDP with discounted cost, there exists an optimal stationary policy [16] π^* such that starting from time t,

$$V_\beta^t(p_t) = V_\beta^{\pi^*,t}(p_t). \tag{7}$$

We only consider the optimal stationary policy because it is tractable and achieves the same minimum cost as any optimal non-stationary policy.

By (5) and (6), the minimum cost function evolves as follows: if an HP coupon is offered at time t, the retailer can perfectly infer the consumer state based on the incurred cost. Therefore,

$$V_{\beta,\mathsf{HP}}^t(p_t) = \beta^t C(p_t, \mathsf{HP}) + (1-p_t)V_\beta^{t+1}(\lambda_{N,A}) + p_t V_\beta^{t+1}(\lambda_{A,A}). \tag{8}$$

If an LP coupon is offered at time t, the retailer cannot infer the consumer state from the cost since both Normal and Alerted consumer impose the same cost C_L. Hence, the discounted cost function can be written as

$$V_{\beta,\mathsf{LP}}^t(p_t) = \beta^t C(p_t, \mathsf{LP}) + V_\beta^{t+1}(p_{t+1}) = \beta^t C_L + V_\beta^{t+1}(T(p_t)). \tag{9}$$

Correspondingly, the minimum cost function is given by

$$V_\beta^t(p_t) = \min\{V_{\beta,\mathsf{LP}}^t(p_t), V_{\beta,\mathsf{HP}}^t(p_t)\}. \tag{10}$$

3 Optimal Stationary Policies

The first main result is a theorem providing the optimal stationary policy for the two-state basic model in Sect. 2.

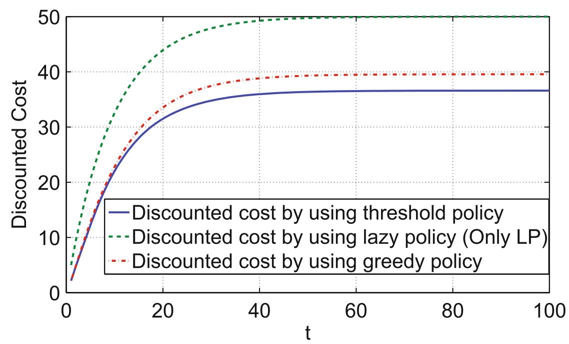


Fig. 2. Discounted cost from by using different decision policies

Theorem 1. *There exists a threshold $\tau \in [0,1]$ such that the following policy is optimal:*

$$\pi^*(p_t) = \begin{cases} \mathsf{LP} & \text{if } \tau \leq p_t \leq 1 \\ \mathsf{HP} & \text{if } 0 \leq p_t \leq \tau \end{cases}. \tag{11}$$

More precisely, assume that $\delta = C_{HA} - C_{HN} + \beta(V_\beta(\lambda_{A,A}) - V(\lambda_{N,A}))$,

$$\tau = \begin{cases} \frac{C_L - (1-\beta)(C_{HN} + \beta V_\beta(\lambda_{N,A}))}{(1-\beta)\delta} & T(\tau) \geq \tau \\ \frac{C_L + \beta\lambda_{N,A}(C_{HA} + \beta V_\beta(\lambda_{A,A})) - (1-\beta(1-\lambda_{N,A}))(C_{HN} + \beta V_\beta(\lambda_{N,A}))}{(1-(\lambda_{A,A} - \lambda_{N,A})\beta)\delta} & T(\tau) < \tau \end{cases}, \tag{12}$$

where for $\lambda_{N,A} \geq \tau$,

$$V_\beta(\lambda_{N,A}) = V_\beta(\lambda_{A,A}) = C_L/(1-\beta) \tag{13}$$

and for $\lambda_{N,A} < \tau$,

$$V_\beta(\lambda_{N,A}) = (1-\lambda_{N,A})[C_{HN} + \beta V_\beta(\lambda_{N,A})] + \lambda_{N,A}[C_{HA} + \beta V_\beta(\lambda_{A,A})], \tag{14}$$

$$V_\beta(\lambda_{A,A}) = \min_{n \geq 0}\{G(n)\}, \tag{15}$$

where

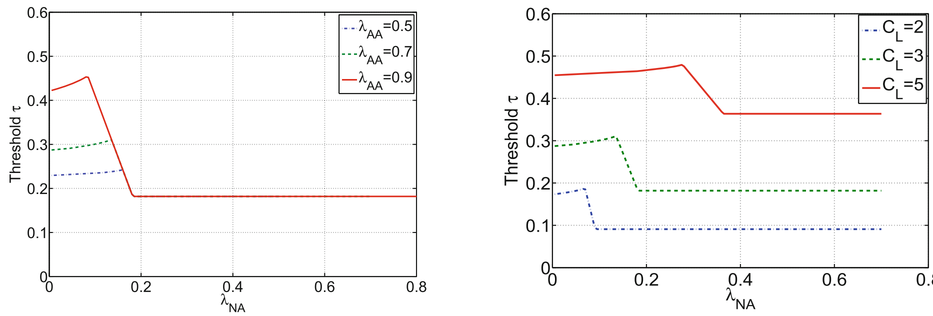
$$G(n) = \frac{C_L \frac{1-\beta^n}{1-\beta} + \beta^n[\bar{T}^n(\lambda_{A,A})(C_{HN} + C(\lambda_{N,A})) + T^n(\lambda_{A,A})C_{HA}]}{1 - \beta^{n+1}[\bar{T}^n(\lambda_{A,A})\frac{\lambda_{N,A}\beta}{1-(1-\lambda_{N,A})\beta} + T^n(\lambda_{A,A})]}, \tag{16}$$

$$T^n(\lambda_{A,A}) = \frac{(\lambda_{A,A} - \lambda_{N,A})^{n+1}(1-\lambda_{A,A}) + \lambda_{N,A}}{1 - (\lambda_{A,A} - \lambda_{N,A})} \tag{17}$$

$$\bar{T}^n(\lambda_{A,A}) = 1 - T^n(\lambda_{A,A}) \tag{18}$$

$$C(\lambda_{N,A}) = \beta\frac{(1-\lambda_{N,A})C_{HN} + \lambda_{N,A}C_{HA}}{1-(1-\lambda_{N,A})\beta}. \tag{19}$$

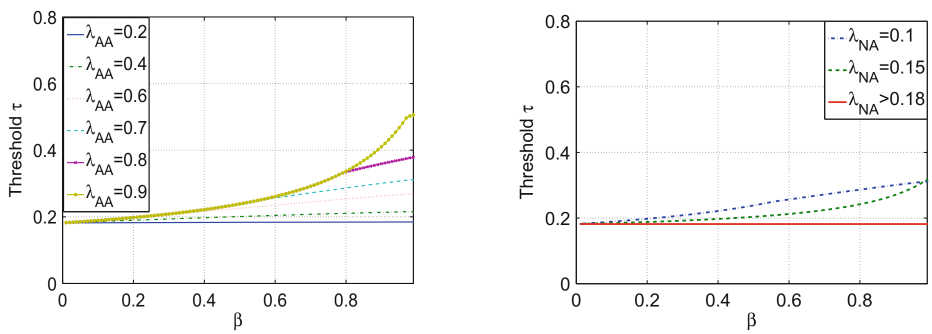
The full proof of Theorem 1 is in the extended version of this paper [17]. We illustrate our policy's performance by comparing its discounted cost to two other



(a) Threshold τ vs. $\lambda_{N,A}$.

(b) Threshold τ vs. $\lambda_{N,A}$.

Fig. 3. Threshold τ vs. β for different values of $\lambda_{A,A}$ and $\lambda_{N,A}$



(a) Threshold τ vs. β for different values of $\lambda_{A,A}$

(b) Threshold τ vs. β for different values of $\lambda_{N,A}$

Fig. 4. Threshold τ vs. β for different values of $\lambda_{A,A}$ and $\lambda_{N,A}$

policies: a greedy policy which minimize the instantaneous cost at each decision epoch and a lazy policy which the retailer only offers LP coupons. Figure 2 shows the discounted cost averaged over 1000 independent MDPs versus the time t for these different decision policies. The illustration demonstrates that the proposed threshold policy performs better than the greedy policy and the lazy policy.

Figure 3a shows the optimal threshold τ as a function of $\lambda_{N,A}$ for three fixed choices of $\lambda_{A,A}$. The threshold increases when $\lambda_{N,A}$ is small because the consumer is less likely to transition from Normal to Alerted so the retailer can more safely offer an HP coupon. When $\lambda_{N,A}$ gets larger, the consumer is more likely to transition from Normal to Alerted, so the retailer is more conservative and decreases the threshold for offering an LP coupon. When $\lambda_{N,A} \geq \kappa$, the retailer uses κ as the threshold for offering an HP coupon. With increasing $\lambda_{A,A}$, the threshold τ decreases. On the other hand, for fixed C_{HN} and C_{HA}, Fig. 3b shows that the threshold τ increases as the cost of offering an LP coupon increases, making it more desirable to take a risk and offer an HP coupon. Figure 4 shows the relationship between the discount factor β and the threshold τ as functions of transition probabilities. Figure 4a shows that τ increases as β increases. When β

is small, the retailer values the present rewards more than future rewards so it is conservative in offering HP coupons to avoid low costs. Figure 4b shows that the threshold is high when $\lambda_{A,A}$ is large or $\lambda_{N,A}$ is small. A high $\lambda_{A,A}$ value indicates that a consumer is more likely to remain in $\mathsf{Alerted}$ state. The retailer is willing to play aggressively since once the consumer is in alerted state, it can take a very long time to transition back to Normal state. A low $\lambda_{N,A}$ value implies that the consumer is not very privacy sensitive. Thus, the retailer tends to offer HP coupons to reduce cost. One can also observe in Fig. 4b that the threshold τ equals to κ after $\lambda_{N,A}$ exceeds the ratio κ. This is consistent with results shown in Fig. 3.

4 Consumer with Multi-level Alerted States

We extend our model to multiple $\mathsf{Alerted}$ states: suppose the consumer state at time t is $G_t \in \{\mathsf{Normal}, \mathsf{Alerted}_1, \dots \mathsf{Alerted}_K\}$, where a consumer in $\mathsf{Alerted}_k$ state is even more cautious about targeted coupons than one in $\mathsf{Alerted}_{k-1}$ state. Define the transition matrix

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{N,N} & \lambda_{N,A_1} & \dots & \lambda_{N,A_K} \\ \lambda_{A_1,N} & \lambda_{A_1,A_1} & \dots & \lambda_{A_1,A_K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{A_K,N} & \lambda_{A_K,A_1} & \dots & \lambda_{A_K,A_K} \end{pmatrix}. \tag{20}$$

We denote $\bar{\mathbf{e}}_i$ to be the i^{th} row of the transition matrix (20). At each time t, the retailer can offer either an HP or an LP coupon. We define $C_{HN}, C_{HA_1}, \dots, C_{HA_K}$ to be the costs of the retailer when an HP coupon is offered while the state of the consumer is Normal, $\mathsf{Alerted}_1, \dots, \mathsf{Alerted}_K$, respectively. If an LP coupon is offered, no matter in which state, the retailer gets a cost of C_L. We assume that $C_{HA_K} \geq \dots \geq C_{HA_1} \geq C_L \geq C_{HN}$. The belief of the consumer being in Normal, $\mathsf{Alerted}_1, \dots, \mathsf{Alerted}_K$ state at time t is defined by $p_{N,t}, p_{A_1,t}, \dots, p_{A_K,t}$, respectively. The expected cost at time t has the following expression:

$$C(\bar{\mathbf{p}}_t, u_t) = \begin{cases} C_L & \text{if } u_t = \mathsf{LP} \\ \bar{\mathbf{p}}_t^T \bar{\mathbf{C}} & \text{if } u_t = \mathsf{HP} \end{cases}, \tag{21}$$

where $\bar{\mathbf{p}}_t = (p_{N,t}, p_{A_1,t}, \dots, p_{A_K,t})^T$ and $\bar{\mathbf{C}} = (C_{HN}, C_{HA_1}, \dots, C_{HA_K})^T$. Assume that the retailer has perfect information about the belief of the consumer state, the cost function evolves as follows: by using an LP coupon at time t,

$$V^t_{\beta,\mathsf{LP}}(\bar{\mathbf{p}}_t) = \beta^t C_L + V^{t+1}_\beta(\bar{\mathbf{p}}_{t+1}) = \beta^t C_L + V^{t+1}_\beta(T(\bar{\mathbf{p}}_t)), \tag{22}$$

where $T(\bar{\mathbf{p}}_t) = \bar{\mathbf{p}}_t^T \mathbf{\Lambda}$ is the one step Markov transition function. By using an HP coupon at time t,

$$V^t_{\beta,\mathsf{HP}}(\bar{\mathbf{p}}_t) = \beta^t \bar{\mathbf{p}}_t^T \bar{\mathbf{C}} + \bar{\mathbf{p}}_t^T \begin{pmatrix} V^{t+1}_\beta(\bar{\mathbf{e}}_1) \\ V^{t+1}_\beta(\bar{\mathbf{e}}_2) \\ \vdots \\ V^{t+1}_\beta(\bar{\mathbf{e}}_{K+1}) \end{pmatrix}. \tag{23}$$

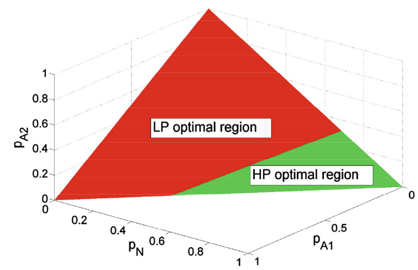


Fig. 5. Optimal policy region for three-state consumer.

Therefore, the minimum cost function is given by (10). In this problem, since the instantaneous costs are nondecreasing with states when the action is fixed and the evolution of belief state is the same for both LP and HP, the existence of an optimal stationary policy with threshold property for finite many states is guaranteed by Proposition 2 in [18]. The optimal stationary policy for a three-state consumer model is illustrated in Fig. 5. For fixed costs, the plot shows the partition of the belief space based on the optimal actions and reveals that offering an HP coupon is optimal when $p_{N,t}$ is high.

5 Consumers with Coupon-Dependent Transition

Generally, consumers' reactions to HP and LP coupons are different. To be more specific, a consumer is likely to feel less comfortable when being offered a coupon on medication (HP) than food (LP). Thus, we assume that the Markov transition probabilities are dependent on the coupon offered. If an LP\HP coupon is offered, the state transition follows the Markov chain

$$\mathbf{\Lambda}_{\mathsf{LP}} = \begin{pmatrix} 1-\lambda_{N,A} & \lambda_{N,A} \\ 1-\lambda_{A,A} & \lambda_{A,A} \end{pmatrix}, \mathbf{\Lambda}_{\mathsf{HP}} = \begin{pmatrix} 1-\lambda'_{N,A} & \lambda'_{N,A} \\ 1-\lambda'_{A,A} & \lambda'_{A,A} \end{pmatrix}, \tag{24}$$

respectively. According to the model in Sect. 2, $\lambda_{A,A} > \lambda_{N,A}, \lambda'_{A,A} > \lambda'_{N,A}$. Moreover, we assume that offering an HP coupon will increase the probability of transition to or staying at Alerted state. Therefore, $\lambda'_{A,A} > \lambda_{A,A}$ and $\lambda'_{N,A} > \lambda_{N,A}$. The minimum cost function evolves as follows:

$$\begin{aligned} V^t_{\beta,\mathsf{HP}}(p_t) &= \beta^t C(p_t, \mathsf{HP}) + (1-p_t)V^{t+1}_\beta(\lambda'_{N,A}) + p_t V^{t+1}_\beta(\lambda'_{A,A}) \\ V^t_{\beta,\mathsf{LP}}(p_t) &= \beta^t C_L + V^{t+1}_\beta(p_{t+1}) = \beta^t C_L + V^{t+1}_\beta(T(p_t)), \end{aligned}$$

where $T(p_t) = \lambda_{N,A}(1-p_t) + \lambda_{A,A}p_t$ is the one step transition defined in Sect. 2.

Theorem 2. *Given action dependent transition matrices* $\mathbf{\Lambda}_{\mathsf{LP}}$ *and* $\mathbf{\Lambda}_{\mathsf{HP}}$*, the optimal stationary policy has threshold structure.*

A full proof of Theorem 2 is in the extended version of this paper [17]. Figure 6 shows the effect of costs on the threshold τ. The threshold for offering an HP

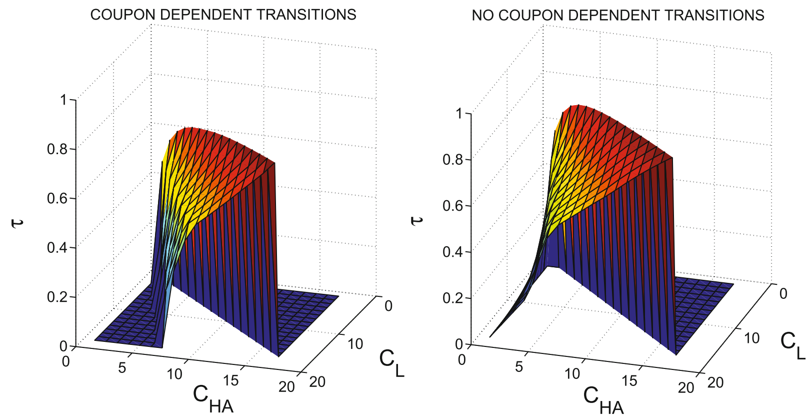


Fig. 6. Optimal τ with/without coupon dependent transition probabilities.

coupon to a consumer with coupon dependent transition probabilities is lower than our original model without coupon-dependent transition probabilities. The retailer can only offer an LP coupon with certain combination of costs; we call this the LP-only region. It can be seen that the LP-only region for the coupon-independent transition case is smaller than that for the coupon-dependent transition case since for the latter, the likelihood of being in an Alerted state is higher for the same costs.

6 Conclusion

We proposed a POMDP model to capture the interactions between a retailer and a privacy-sensitive consumer in the context of personalized shopping. The retailer seeks to minimize the expected discounted cost of violating the consumer's privacy. We showed that the optimal coupon-offering policy is a stationary policy that takes the form of an explicit threshold that depends on the model parameters. In summary, the retailer offers an HP coupon when the Normal to Alerted transition probability is low or the probability of staying in Alerted state is high. Furthermore, the threshold optimal policy also holds for consumers whose privacy sensitivity can be captured via multiple alerted states as well as for the case in which consumers exhibit coupon-dependent transition. Our work suggests several interesting directions for future work: cases where retailer has additional uncertainty about the state, for example due to randomness in the received costs, game theoretic models to study the interaction between the retailer and strategic consumers, and more generally, understanding the tension between acquiring information about the consumers and maximizing revenue.

References

1. Hill, K.: How target figured out a teen girl was pregnant before her father did (2012). http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-outa-teen-girl-was-pregnant-before-her-father-did/

2. Dwork, C.: Differential privacy. In: van Tilborg, H.C.A., Jajodia, S. (eds.) Encyclopedia of Cryptography and Security, pp. 338–340. Springer, New York (2011)
3. Sankar, L., Kar, S., Tandon, R., Poor, H.V.: Competitive privacy in the smart grid: an information-theoretic approach. In: 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 220–225. IEEE (2011)
4. Acquisti, A.: The economics of personal data and the economics of privacy. Background Paper for OECD Joint WPISP-WPIE Roundtable, vol. 1 (2010)
5. Ghosh, A., Roth, A.: Selling privacy at auction. Games Econ. Behav. (2013). Elsevier
6. Hsu, J., Huang, Z., Roth, A., Roughgarden, T., Wu, Z.S.: Private matchings and allocations. arXiv preprint arXiv:1311.2828 (2013)
7. Venkitasubramaniam, P.: Privacy in stochastic control: a markov decision process perspective. In: Proceedings of Allerton Conference, pp. 381–388 (2013)
8. Feinberg, E.A., Shwartz, A., Altman, E.: Handbook of Markov Decision Processes: Methods and Applications. Kluwer Academic Publishers, Boston (2002)
9. Lipsa, G.M., Martins, N.C.: Remote state estimation with communication costs for first-order LTI systems. IEEE Trans. Autom. Control **56**(9), 2013–2025 (2011)
10. Nayyar, A., Başar, T., Teneketzis, D., Veeravalli, V.V.: Optimal strategies for communication and remote estimation with an energy harvesting sensor. IEEE Trans. Autom. Control **58**(9), 2246–2260 (2013)
11. MacPhee, I., Jordan, B.: Optimal search for a moving target. Probab. Eng. Informational Sci. **9**(02), 159–182 (1995)
12. Mansourifard, P., Javidi, T., Krishnamachariy, B.: Tracking of real-valued continuous markovian random processes with asymmetric cost and observation. In: American Control Conference (2015)
13. Ross, S.M.: Quality control under markovian deterioration. Manag. Sci. **17**(9), 587–596 (1971)
14. Laourine, A., Tong, L.: Betting on gilbert-elliot channels. IEEE Trans. Wirel. Commun. **9**(2), 723–733 (2010)
15. Bertsekas, D.P.: Dynamic Programming and Optimal Control, vol. 1, 2, issue 2. Athena Scientific, Belmont (1995)
16. Ross, S.M.: Applied Probability Models with Optimization Applications. Courier Dover Publications, New York (2013)
17. Huang, C., Sankar, L., Sarwate, A.D.: Designing incentive schemes for privacy-sensitive users. ArXiV, Technical report arXiv:1508.01818 [cs.GT], August 2015
18. Lovejoy, W.S.: Some monotonicity results for partially observed markov decision processes. Oper. Res. **35**(5), 736–743 (1987)