Private Filtering for Hidden Markov Models

Rami Mochaourab and Tobias J. Oechtering

Abstract—Consider a hidden Markov model describing a system with two types of states: a monitored state and a private state. The two types of states are dependent and evolve jointly according to a Markov process with a stationary transition probability. It is desired to reveal the monitored states to a receiver but hide the private states. For this purpose, a privacy filter is necessary which suitably perturbs the monitored states before communication to the receiver. Our objective is to design the privacy filter to optimize the trade-off between monitoring accuracy and privacy, measured through a time-invariant distortion measure and Shannon’s equivocation, respectively. As the optimal privacy filter is difficult to compute using dynamic programming, we adopt a suboptimal greedy approach through which the privacy filter can be computed efficiently. Here, the greedy approach has the additional advantage of not being restricted to finite time horizon setups. Simulations show the superiority of the approach compared to a privacy filter which only adds independent noise to the observations.

Index Terms—Hidden Markov models; Privacy; Dynamic programming; Greedy algorithm

I. INTRODUCTION

As we increasingly rely on Internet services that require sensitive information from us, such as those related to monitoring of health [1], human activity [2], or smart homes [3], privacy becomes a crucial issue since the information we provide could reveal additional information which is private. Storing and analysis of large amounts of data is now becoming practice as memory becomes cheaper and processing tools becoming more powerful. For this reason, it is necessary to design suitable mechanisms which can preserve privacy while maintain Internet service benefits.

Several privacy notions and mechanisms have been proposed and studied in the literature. Considering static databases, one recent popular privacy notion is that of differential privacy [4]. The mechanism perturbs the data in such a way that statistical results from queries to the database remain useful but it becomes difficult to reveal information about a specific entry in the database. Privacy in static databases has been also studied using information theoretic tools in [5], [6]. The associated privacy-preserving mechanisms introduce uncertainty about the private information by constructing distorted versions of the communicated messages. It is shown in [5], [6] that characterizing the trade-off between the accuracy of an answer to a query and the privacy guarantees is related to rate-distortion theory, and convex optimization techniques can be used to characterize this trade-off.

In contrast to static databases, privacy for dynamical systems, such as cyber-physical systems, requires more sophisticated measures due to the additional time dimension [7]–[9]. In [7], differential privacy is defined and analyzed for general time-varying data streams. The objective is to generate accurate statistics over the individuals’ data, while guaranteeing certain level of uncertainty about the contribution from an individual. In [8], [9], stochastic control problems are considered and efficient decision policies are computed for controlling the system with the privacy requirement of minimizing the information leakage about the system’s states.

In this paper, we consider hidden Markov models which are known to describe several interesting applications [10], such as human activities [2] or machine states [11]. We assume that the considered system possesses two types of hidden states, X-states and Z-states, that evolve jointly according to a Markov process. The two types of states are dependent, as for example a human activity (X-state) of going up the stairs or running and a private medical condition or mobility impairment (Z-state) which can be inferred from monitoring this activity. It is desired to estimate the sequences of X-states as accurately as possible at a receiver but keep the sensitive information (Z-states) as private as possible. Our model differs from those in [7]–[9] through considering such two types of state sequences with applications in monitoring problems.

The problem we consider in this paper is that of minimizing the weighted sum of two objectives: the estimation cost for the X-states and the privacy cost for the Z-states, which is based on Shannon’s equivocation. First, we formulate a suitable upper bound on the privacy cost that can be optimized efficiently. Afterwards, we analyze the optimal privacy filter through dynamic programming and show that the solution is still fairly difficult to compute. Consequently, we resort to a greedy approach which delivers an upper bound on the solution of the original problem. The performance of the proposed privacy filter is shown through simulations to be superior to a more general privacy filter design which only adds independent noise to the observations.

Notations: Let \( \Delta(\mathcal{A}) \) be the set of all probability distributions over the set \( \mathcal{A} \). We define the collection of variables as \( X_{1:T} = \{X_1, \ldots, X_T\} \).
II. SYSTEM MODEL

Consider a finite state hidden Markov model [10], as illustrated in Fig. 1. The state space for the hidden variables is $X \times Z$, where $X$ and $Z$ are finite sets. At a time $t$, the hidden states are described by the random variables $X_t \in X$ and $Z_t \in Z$, where $X_t$ is the state that should be delivered to the receiver with good accuracy while $Z_t$ is the private state which should be kept hidden as much as possible.

Initially, we will assume a finite time horizon with stages $t = 0, 1, \ldots, T$. We will drop this assumption in Section IV-B when we consider greedy optimization of the privacy filter. The sequence of states $(X_0, Z_0), (X_1, Z_1), \ldots, (X_T, Z_T)$ forms a Markov chain and evolve according to the following stationary transition probability distribution

$$
\theta(x_t, z_t, x_{t-1}, z_{t-1}) \triangleq \mathbb{P}\{X_t = x_t, Z_t = z_t | X_{t-1} = x_{t-1}, Z_{t-1} = z_{t-1}\}. \tag{1}
$$

Let $\pi_0$ be the distribution of the initial states $(X_0, Z_0)$.

A. Privacy Filter

We assume that the privacy filter, as is illustrated in Fig. 1, is memoryless, i.e., it takes the current states $(x_t, z_t)$ as input to generate $y_t \in Y$ according to the conditional distribution

$$
\phi_t(y_t, x_t, z_t) \triangleq \mathbb{P}\{Y_t = y_t | X_t = x_t, Z_t = z_t\}. \tag{2}
$$

We will refer to $\phi_t : X \times Z \rightarrow \Delta(Y)$ as the privacy filter at time $t$. The assumption that the privacy filter is memoryless is mainly motivated for supporting low complexity mechanisms. Observe that the privacy filter is time-dependent and will be optimized later using all available model information.

B. Receiver

The receiver uses a given and fixed function $f : Y \rightarrow X$ to map each output $y_t$ to an estimate of $x_t$. The accuracy of this estimate thus depends on the privacy filter, and one of the objectives in the optimization of the privacy filter will be to minimize the distortion in the estimates.

In general, by knowing the $X$-states at the receiver, information about the $Z$-states can be inferred. We assume that the receiver is “curious” [6] in the sense that it is interested in knowing the $Z$-states. Similar to [6], we assume “worst-case” statistical side information at the receiver, such that the receiver knows the privacy filter $\phi_{1:T}$, the transition probability distribution $\theta$, and the distribution of the initial states of the system $\pi_0$. With this information and the outputs $y_{1:t}$, the receiver is able to generate estimates for the sequence of states using Bayes' estimation [12, ch. 3].

At a time $t$, the receiver uses Bayes’ rule to calculate the posterior probability of $(X_t, Z_t)$ given the outputs $y_{1:t}$

$$
\pi_t(x, z) \triangleq \mathbb{P}\{X_t = x, Z_t = z | Y_{1:t} = y_{1:t}\} = \frac{\phi_t(y_t, x, z)\psi_t(x, z, \pi_{t-1})}{\gamma_t(y_t, \phi_t, \pi_{t-1})}, \tag{3}
$$

where $\phi_t(y_t, x, z)$ is the privacy filter at time $t$, defined in (2), $\psi_t(x, z, \pi_{t-1})$ is the prior for $(X_t, Z_t)$

$$
\psi_t(x, z, \pi_{t-1}) \triangleq \mathbb{P}\{X_t = x, Z_t = z | Y_{1:t-1} = y_{1:t-1}\} = \sum_{x' \in X} \sum_{z' \in Z} \theta(x', z, z', \pi_{t-1}(x', z')), \tag{4}
$$

and $\gamma_t(y_t, \phi_t, \pi_{t-1})$ is the likelihood of $y_t$

$$
\gamma_t(y_t, \phi_t, \pi_{t-1}) \triangleq \mathbb{P}\{Y_t = y_t | Y_{1:t-1} = y_{1:t-1}\} = \sum_{x \in X} \sum_{z \in Z} \phi_t(y_t, x, z)\psi_t(x, z, \pi_{t-1}). \tag{5}
$$

Note that the posterior in (3) depends on the outputs $y_{1:t}$ generated according to $\phi_{1:t}$. Alternatively, using its recursive structure, $\pi_t$ depends on the privacy filter $\phi_t$, the posterior of the previous time step $\pi_{t-1}$, and the current output $y_t$ only. Here, $\pi_{t-1}$ is called information state [13], since $\pi_{t-1}$ is a function of $y_{1:t-1}$, and $\pi_t$ can be determined from $\pi_{t-1}$ and $y_t$ given $\phi_t$.

III. ACCURACY AND PRIVACY COSTS

The objective is to design the privacy filter $\phi_{1:T}$ such that the estimates $f(y_1), \ldots, f(y_T)$ of the $X$-states are as accurate as possible, while $Z_{1:T}$ are kept as private as possible. If $X_{1:T}$ and $Z_{1:T}$ are dependent, then the objectives are conflicting and our goal is to design a filter with an optimality trade-off. To this end, we first define the measures we will use to quantify each of the objectives and then we will state the optimization problem using a weighted sum of the two objectives.

At a time $t$, let $F(f(y_t), x_t)$ be a distortion measure between the state $x_t$ and its estimate $f(y_t)$. For example, $F$ can be the Euclidean distance $F(f(y_t), x_t) = ||f(y_t) - x_t||_2$. The accuracy cost at time $t$ is defined as the expected distortion $[5]$

$$
d_t(\pi_{t-1}, \phi_t) \triangleq \mathbb{E}[F(f(Y_t), X_t)] = \sum_{y \in Y} \sum_{x \in X} \sum_{z \in Z} \phi_t(y, x, z)\psi_t(x, z, \pi_{t-1})F(f(y), x). \tag{6}
$$

Consequently, we can define the total cost of estimating the sequence of $X$-states for a given privacy filter $\phi_{1:T}$ as

$$
D(\phi_{1:T}) \triangleq \frac{1}{T} \sum_{t=1}^{T} d_t(\pi_{t-1}, \phi_t). \tag{7}
$$

In order to suitably quantify the privacy cost, we need a measure which captures the uncertainty in the random variables $Z_{1:T}$. A suitable measure is Shannon’s equivocation (see discussion in [8]) which we will use as privacy cost

$$
C(\phi_{1:T}) \triangleq \log |Z| - \frac{1}{T} H(Z_{1:T} | Y_{1:T}). \tag{9}
$$

For the conditional entropy $H(Z_{1:T} | Y_{1:T})$, the distribution of the $Z$-states correspond to the posterior defined in (3). Observe, that the privacy cost $C(\phi_{1:T})$ attains its minimum when the uncertainty about the $Z$-states is largest, and reaches its maximum if the $Z$-States can be predicted perfectly.

Using the chain rule for conditional entropy [14], we have

$$
C(\phi_{1:T}) = \frac{1}{T} \sum_{t=1}^{T} c_t(\phi_{1:t}), \tag{10}
$$

where

$$
c_t(\phi_{1:t}) \triangleq \log |Z| - H(Z_t, Y_t | Z_{1:t-1}, Y_{1:t-1}) + H(Y_t | Y_{1:t-1}). \tag{11}
$$

The factorization above reveals a causal structure which is convenient for the optimization of the privacy filter. However, since the privacy filter and the transition probability of the $Z$-states depend on the $X$-states as well (not present in (11)), we will need to use the following bound on the privacy cost.
Lemma 1: At time $t$, an upper bound for the immediate privacy cost $c_t(\phi_{1:T})$ in (11) is
\begin{align}
\hat{c}_t(\pi_{t-1}, \phi_t) \triangleq & \log |Z| - H(Z_t|X_{t-1}, Z_{t-1}) \\
& - H(Y_t|X_t, Z_t) + H(Y_t|Y_{1:t-1}) \quad (12)
\end{align}

Proof: The proof is provided in Appendix A.

In (12), the conditional entropy $H(Z_t|X_{t-1}, Z_{t-1})$ corresponds to the uncertainty due to the transition probability $\theta$, given in (1). The entropy terms $H(Y_t|X_t, Z_t)$ and $H(Y_t|Y_{1:t-1})$ depend on the privacy filter $\phi_t$ and can be computed utilizing the knowledge of $\pi_{t-1}$. While $H(Y_t|X_t, Z_t)$ lowers the privacy cost when more randomization exists in the privacy filter, the term $H(Y_t|Y_{1:t-1})$ increases the privacy cost whenever $Y_t$ holds information not present in $Y_{1:t-1}$.

Using Lemma 1, we obtain
\begin{align}
\hat{C}(\phi_{1:T}) \triangleq & \frac{1}{T} \sum_{t=1}^{T} \hat{c}_t(\pi_{t-1}, \phi_t) \quad (13)
\end{align}
as an upper bound on the privacy cost in (9).

IV. PRIVACY FILTER DESIGN

Using the accuracy and privacy costs in (8) and (13), respectively, we formulate the privacy filter optimization problem as
\begin{align}
\phi_{1:T}^* = \arg \min_{\phi_{1:T}} \lambda D(\phi_{1:T}) + (1 - \lambda) \hat{C}(\phi_{1:T}) \quad (14)
\end{align}
with $\phi_t: \mathcal{X} \times \mathcal{Z} \rightarrow \Delta(Y)$, $t = 1, \ldots, T$ and $\lambda \in [0, 1]$. We can rewrite (14), using (8) and (13), as follows
\begin{align}
\phi_{1:T}^* = \arg \min_{\phi_{1:T}} \frac{1}{T} \sum_{t=1}^{T} g_t(\pi_{t-1}, \phi_t) \quad (15)
\end{align}
where
\begin{align}
g_t(\pi_{t-1}, \phi_t) \triangleq & \lambda d_t(\pi_{t-1}, \phi_t) + (1 - \lambda) \hat{c}_t(\pi_{t-1}, \phi_t) \quad (16)
\end{align}
The additive structure in the objective function above makes the optimization using dynamic programming possible.

A. Dynamic Programming Approach

Problem (15) can be solved by dynamic programming using Bellman optimality equations [13], [15]. The dynamic programming algorithm generates a sequence of value functions $V_{\tau}: \Delta(\mathcal{X} \times \mathcal{Z}) \rightarrow \mathbb{R}$ with $\tau = 0, \ldots, T$ corresponding to the number of stages to go until the final stage $T$ (i.e., associated stage index $t = T - \tau$). The value function with $\tau$ stages remaining can be formulated using the transformation in [15]
\begin{align}
V_{\tau}(\beta) = & \max_{\phi_t} \left\{ g_t(\beta, \phi_t) + \sum_{y \in \mathcal{Y}} \gamma_t(y, \phi_t, \beta)V_{\tau-1}(\pi_t) \right\} \quad (17)
\end{align}
where $\beta \in \Delta(\mathcal{X} \times \mathcal{Z})$ represents a continuous state, corresponding to the belief about the hidden states, and $\pi_t$ is defined in (3). The algorithm, known as backward induction, initializes with $\tau = 0$ and $V_0(\beta) = 0$ for all $\beta \in \Delta(\mathcal{X} \times \mathcal{Z})$. For each increment of $\tau$ until $\tau = T$, $V_{\tau}(\beta)$ is optimized for every $\beta \in \Delta(\mathcal{X} \times \mathcal{Z})$.

The difficulty in the dynamic programming approach is due to the complexity of efficiently calculating, representing, and storing the value functions $V_{\tau}(\beta)$, which are functions of continuous variables (the belief $\beta$) [16]. Similar to [8], we take the approach of optimizing an upper bound of our objective which is efficiently computable through a greedy approach.

B. Greedy Approach

The greedy approach optimizes the immediate costs at each time $t$ utilizing the posterior probability $\pi_{t-1}$ calculated at the previous time step. The greedy algorithm is as follows

1) Initialize: $t = 1$
2) Calculate $\phi_t^* = \arg \min_{\phi_t} g_t(\pi_{t-1}, \phi_t)$ and generate $y_t$
3) Calculate $\pi_t$ using (3)
4) Stop if $t = T$. Otherwise, $t = t + 1$ and go to Step 2.

The calculation of the privacy filter in Step 2 can be performed efficiently as shown in the next result.

Theorem 1: The problem minimize$_{\phi_t} g_t(\pi_{t-1}, \phi_t)$ can be formulated as a convex program.

Proof: The proof is provided in Appendix B.

Notice that in the greedy algorithm, it is also possible to remove the stopping criterion in Step 4. That is, the greedy approach is not limited to the finite time interval $T$.

V. SIMULATIONS

For the simulations, we consider binary state spaces $\mathcal{X} = \mathcal{Z} = \{0, 1\}$ and time horizon $T = 250$. The initial probability distribution $\pi_0$ is uniform over $\mathcal{X} \times \mathcal{Z}$. We assume that the states satisfy the Markov chain $Z_{t-1} \leftrightarrow X_{t-1} \leftrightarrow X_t \leftrightarrow Z_t$. For example, this assumption could relate to applications in which the $Z$-states correspond to health conditions of a person or a machine, and the $X$-states are associated treatments. The private health state at a time $t$ depends on the immediate treatment and vice versa. For this model, the transition probability distribution in (1) reduces to
\begin{align}
\theta(x, z, x', z') = & \mathbb{P}\{Z_t = z|X_t = x, X_{t-1} = x', Z_{t-1} = z'\} \\
& \cdot \mathbb{P}\{X_t = x|X_{t-1} = x', Z_{t-1} = z'\} = \mathbb{P}\{Z_t = z|x_t = x\} \mathbb{P}\{X_t = x|X_{t-1} = x'\}. \quad (18)
\end{align}
We use binary symmetric channels to specify the transition probabilities in (18) as follows:
\begin{align}
\mathbb{P}\{X_t = x_t|X_{t-1} = x_{t-1}\} = & \begin{cases} 1 - \delta & x_t = x_{t-1} \\ \delta & \text{otherwise} \end{cases}, \quad (19)
\end{align}
\begin{align}
\mathbb{P}\{Z_t = z_t|x_t = x_t\} = & \begin{cases} 1 - \epsilon & z_t = x_t \\ \epsilon & \text{otherwise} \end{cases}, \quad (20)
\end{align}
where $\delta, \epsilon \in [0, 1]$. Here, for given $x_t, \epsilon$ affects the uncertainty in the $Z$-states at the receiver. In the extreme case, for $\epsilon = 0.5$, the $Z$-states are independent of the $X$-states.

We assume that the output set is binary, $\mathcal{Y} = \{0, 1\}$, and at a time $t$, the output $y_t$ corresponds to the estimate of $x_t$. That is, the function $f$, defined in Section II-B, is the identity mapping $f(y) = y$. We use the Hamming distortion measure between the estimate $y_t$ and the actual state $x_t$ as $F(y_t, x_t) = |y_t - x_t|$, and the immediate accuracy cost at a time $t$ corresponds to the average distortion as defined in (6).

Fig. 2 shows the trade-off between the privacy and accuracy costs achieved by the greedy algorithm of Section IV-B. The trade-off, achieved by executing the greedy algorithm for different values of $\lambda \in [0, 1]$, corresponds to an upper bound to the optimal trade-off between privacy and accuracy. This is due to the upper bound for the privacy cost in Lemma 1 and
optimization. This issue has been also argued in [6].

The trade-off between privacy and accuracy using the prior independent privacy filter which adds independent noise to the observations, such as that of differentially private filtering [7]. This filter can be specified for our setup as follows

$$
\phi_t^{\text{in}}(y_t, x_t, z_t) = \begin{cases} 
\eta & y_t = x_t \\
1 - \eta & \text{otherwise}
\end{cases}
$$

where $\eta \in [0, 1]$. For $\eta = 0.5$, full privacy is ensured which is equivalent to the optimization of our privacy filter with $\lambda = 0$. For $\eta = 1$, largest accuracy is achieved at the cost of privacy, corresponding to $\lambda = 1$ in our privacy filter optimization. The trade-off between privacy and accuracy using the prior independent privacy filter, by varying $\eta \in [0, 0.5, 1]$ in (21), is shown to be inferior to our privacy filter. This reveals that it is necessary to take the prior into account in the privacy filter optimization. This issue has been also argued in [6].

In Fig. 3, we plot the upper bound on the privacy cost for different $\epsilon$ values. For $\epsilon = 0$, the $Z$-states are fully determined by the $X$-states and hence the privacy cost is largest. For $\epsilon = 0.5$, the $Z$-states and the $X$-states are independent. In this case, the privacy cost is lowest. Clearly, the actual privacy cost for $\epsilon = 0.5$ should be zero, but the gap is due to the upper bound.

VI. CONCLUSIONS

We have studied a privacy mechanism in a setup in which it is desired to reveal monitored states to a receiver but hide related private states. Here, the monitored and private states evolve jointly according to a stationary Markov transition probability. The privacy filter is efficiently found by a greedy optimization approach which has the objective of characterizing the trade-off between monitoring accuracy and privacy. Simulations illustrate this trade-off and show the superiority of the approach compared to a prior independent privacy filter.

APPENDIX

A. Proof of Lemma 1

An upper bound for the second term in (11) is

$$
-H(Z_t, Y_t | Z_{1:t-1}, Y_{1:t-1}) = -H(Y_t | Z_{1:t-1}, Y_{1:t-1}) - H(Y_t | Z_{1:t}, Y_{1:t-1})
\leq -H(Y_t | X_{t-1}, Z_{1:t-1}, Y_{1:t-1}) - H(Y_t | X_t, Z_{1:t-1}, Y_{1:t-1}),
$$

where the inequality is due to conditioning which reduces the entropy. Using the Markov property, we have

$$
H(Z_t | X_{t-1}, Z_{1:t-1}, Y_{1:t-1}) = H(Z_t | X_{t-1}, Z_{1:t-1}).
$$

Since the privacy filter at time $t$, defined in (2), depends on the current states $X_t$ and $Z_t$ only, we have

$$
H(Y_t | X_t, Z_{1:t}, Y_{1:t-1}) = H(Y_t | X_t, Z_t).
$$

B. Proof of Theorem 1

The function $g_t(\pi_{t-1}, \phi_t)$ in (16) is a weighted sum of $d_t(\pi_{t-1}, \phi_t)$ and $\tilde{c}_t(\pi_{t-1}, \phi_t)$. The affine function $d_t(\pi_{t-1}, \phi_t)$ in (6) is convex in $\phi_t(y_t, x_t, z_t)$ for $x_t, z_t, y_t \in X \times Z \times Y$. Next, we show that the privacy cost $\tilde{c}_t(\pi_{t-1}, \phi_t)$ in (12) is a convex function. Its first entropy term

$$
-H(Z_t | X_{t-1}, Z_{1:t-1}) = \sum_{x' \in X} \sum_{z' \in Z} \pi_{t-1}(x', z') \log \sum_{x' \in X} \theta(x, z, x', z')
$$

only depends on the transition probability distribution $\theta$ and $\pi_{t-1}$. We can expand the second entropy term in (12) as

$$
-H(Y_t | X_t, Z_t) = \sum_{x \in X} \sum_{z \in Z} \sum_{y \in Y} \Pr(Y_t = y, X_t = x, Z_t = z) \log \phi_t(y, x, z)
= \sum_{x \in X} \sum_{z \in Z} \sum_{y \in Y} \psi_t(x, z, \pi_{t-1}) \phi_t(y, x, z) \log \phi_t(y, x, z),
$$

and the third entropy term in (12) as

$$
H(Y_t | X_t, Z_t) = -\sum_{y \in Y} \gamma_t(y, \phi_t, \pi_{t-1}) \log \gamma_t(y, \phi_t, \pi_{t-1})
= -\sum_{x \in X} \sum_{z \in Z} \sum_{y \in Y} \psi_t(x, z, \pi_{t-1}) \phi_t(y, x, z) \log \gamma_t(y, \phi_t, \pi_{t-1}).
$$

Then,

$$
-H(Y_t | X_t, Z_t) + H(Y_t | Y_{1:t-1}) = \sum_{x \in X} \sum_{z \in Z} \sum_{y \in Y} \psi_t(x, z, \pi_{t-1}) \phi_t(y, x, z) \log \frac{\phi_t(y, x, z)}{\gamma_t(y, \phi_t, \pi_{t-1})}.
$$

Equation (24) is a weighted sum of functions of the form $G(v, x) = v^T \log \frac{v}{x}$ with $v > 0$. $G(v, x)$ is the perspective [17, ch. 3.2.6] of the convex function $x \log x$ which is jointly convex in $v$ and $x$. Consequently, (24) is jointly convex in $\phi_t(y, x, z)$ and $\gamma_t(y, \phi_t, \pi_{t-1})$ for $x, z, y \in X \times Z \times Y$.

It follows that $g_t(\pi_{t-1}, \phi_t)$ is jointly convex in $\phi_t$ and $\gamma_t$. Then we can optimize $g_t(\pi_{t-1}, \phi_t)$ w.r.t $\phi_t$ and $\gamma_t$ utilizing the following coupling constraint for all $y \in Y$

$$
\gamma_t(y, \phi_t, \pi_{t-1}) = \sum_{x \in X} \sum_{z \in Z} \psi_t(x, z, \pi_{t-1}) \phi_t(y, x, z).
$$
REFERENCES


