Preserving ON-OFF Privacy for Past and Future Requests

Fangwei Ye, Carolina Naim, Salim El Rouayheb
Department of Electrical and Computer Engineering, Rutgers University
Emails: {fangwei.ye, carolina.naim, salim.elrouayheb}@rutgers.edu

***Abstract*—We study the ON-OFF privacy problem. At each time, the user is interested in the latest message of one of N sources. Moreover, the user is assumed to be incentivized to turn privacy ON or OFF whether he/she needs it or not. When privacy is ON, the user wants to keep private which source he/she is interested in. The challenge here is that the user's behavior is correlated over time. Therefore, the user cannot simply ignore privacy when privacy is OFF, because this may leak information about his/her behavior when privacy was ON due to correlation.**

We model the user's requests by a Markov chain. The goal is to design ON-OFF privacy schemes with optimal download rate that ensure privacy for past and future requests. The user is assumed to know future requests within a window of positive size ω and uses it to construct privacy-preserving queries. In this paper, we construct ON-OFF privacy schemes for $N = 2$ sources and prove their optimality.

I. Introduction

Privacy of online users has become a major concern. Without agreeing to it, users unknowingly leak valuable personal information, such as sex, age, health disorders, political views, etc., through their daily online activities. Several existing privacy-preserving solutions can be utilized to ensure a desired level of privacy for the user, such as anonymity [1], differential privacy [2], private information retrieval [3], to name a few.

In all the privacy problems above, it is assumed that the user always wants to be private. Privacy, however, is expensive. Privacy-preserving protocols incur higher computational costs on the service provider, and typically lead to degraded quality of service and larger delays at the user side [4].

This motivates us to think of privacy as an expensive utility, which should be turned OFF when not needed. Much like one turns off the lights before leaving home. The user may want to turn his/her privacy ON or OFF depending on the internet connection he/she is using, his/her location or his/her device used to get online, etc. This behavior of the user may be incentivized by the service providers who encourage him/her to require privacy only when it is needed.

The challenge in designing algorithms that enable privacy to be switched between ON and OFF, and vice versa, is that the user's behavior is correlated over time. This is essentially true because the user's choices are personal and are not independent over time. For instance, a user watching online videos, will most likely pick the next video to watch from a suggested personalized list that is specifically curated for him/her. Therefore, the user cannot simply ignore privacy when privacy is OFF, because this may leak information about his/her behavior when privacy was ON due to correlation.

To capture this challenge, the authors introduced the ON-OFF privacy problem in [5]. A user is interested in the latest message generated by one of N sources. Think, for example, a user is subscribed to $N = 2$ political YouTube channels, one is pro-right and one is pro-left. Occasionally, the user wants to watch the latest video on one of these channels. He/she has a choice between turning privacy ON or OFF. When privacy is ON, the user is not interested in hiding which particular video he/she wants to watch. Rather, he/she is interested in hiding the channel on which that video is posted, because he/she does not want to reveal his/her political interests. In general, when privacy is ON, the user wants to hide which message of the N sources he/she is interested in.

In [5], we studied ON-OFF privacy in which it was required to ensure privacy for past requests for which privacy was turned ON. In this paper, we consider a more stringent privacy requirement and want to preserve privacy for both *past and future* requests. We follow a setup similar to the one in [5] in which the user's request are modeled by a Markov chain, but with one significant difference. We assume here that the user knows the requests in a small window of positive size $\omega > 0$ in the future. In practice, this may happen in applications where the user can queue up his/her requests, such as when watching online videos.

Under this new setting, we study the download rate, which is measured by the ratio of the average length of downloaded data to the message length. We characterize the optimal download rate for the system with $N = 2$ sources and provide explicit constructions of ON-OFF privacy schemes that achieve it. One interesting implication of our result, is that the optimal rate does not depend on the window size. Thus, a window of size $\omega = 1$ is sufficient to achieve the optimal rate.

II. Problem Formulation

There is a single server storing N information sources indexed by $\mathcal{N} := \{1, \ldots, N\}$. Each source generates an independent message $W_{x,t}$ at time t, where $x \in \mathcal{N}$. We assume that $t \in \mathbb{N}$ throughout this paper.

A user is interested in one of the sources at each time, and wishes to retrieve the latest message generated by the corresponding source. In particular, let X_t be the index of the

This work was supported by NSF Grant CCF 1817635.

978-1-5386-6900-6/19/$31.00 ©2019 IEEE

desired source at time t, and in the sequel we call X_t the user's request. By slightly abusing the notation, we denote the latest message generated by the desired source X_t by $W_{X_t,t}$, and the user wishes to retrieve the message $W_{X_t,t}$. We assume that the messages $\{W_{x,t} : x \in \mathcal{N}, t \in \mathbb{N}\}$ are mutually independent, and each message consists of L symbols. Without loss of generality, we assume that each of the messages is uniformly distributed over $\{0,1\}^L$, *i.e.,* $H(W_{x,t}) = L$, and

$$H(W_{x,t} : x \in \mathcal{N}, t \in \mathbb{N}) = \sum_{x,t} H(W_{x,t}). \qquad (1)$$

The user's requests are generated by a discrete-time information source $\{X_t : t \geq 0\}$. In this paper, we are particularly interested in the case where the requests $\{X_t : t \geq 0\}$ are Markov. The transition matrix M of the Markov chain is assumed to be known by both the server and the user.

At time t, the user may or may not wish to keep the identity of the source being interested in. Specifically, the *privacy mode* F_t at time t can be either ON or OFF, where F_t is ON when the user wishes to keep X_t private, while F_t is OFF when the user is not concerned with privacy. The privacy mode is also assumed to be known by the server.

The user is allowed to generate unlimited local randomness, and we are not interested in the amount of randomness used. Therefore, we assume without loss of generality that the random variables $\{S_t : t \geq 0\}$, representing the local randomness, are mutually independent.

All information sources are assumed to be independent, that is, the user's requests $\{X_t : t \geq 0\}$, the privacy mode $\{F_t : t \geq 0\}$, the messages $\{W_{x,t} : x \in \mathcal{N}, t \geq 0\}$ and the local randomness $\{S_t : t \geq 0\}$ are mutually independent.

At time t, the user will construct a query Q_t and send it to the server. Upon receiving the query, the server responds by producing an answer Y_t. After receiving the answer, the user should be able to decode $W_{X_t,t}$ correctly.

We assume that the user knows the future requests in a window of positive size ω. This means at time t, the user knows the future requests $\{X_{t+1}, \ldots, X_{t+\omega}\}$ in addition to all past (including current) requests $\{X_0, \ldots, X_t\}$. In practice, it often happens that the user has some side information to predict his/her requests in the near future. Later, we will show that only a window of size $\omega = 1$ is needed.

The query Q_t at time t is generated by the query encoding function ϕ_t, which is assumed to be a function of the *causal* information, *i.e.,* previous requests and local randomness $\{X_i, S_i : i \leq t\}$, and future requests $\{X_{t+1}, \ldots, X_{t+\omega}\}$ for some $\omega \in \mathbb{N}$. Hence, we assume that

$$Q_t = \phi_t\left(X_{[t+\omega]}, S_{[t]}\right), \qquad (2)$$

where $[t+\omega] := \{0, 1, \ldots, t+\omega\}$.

Accordingly, the answer Y_t of the server is given by the answer encoding function ρ_t, which is assumed to be a function of the query Q_t and the latest messages, *i.e.,*

$$Y_t = \rho_t\left(Q_t, W_{1,t}, \ldots, W_{N,t}\right). \qquad (3)$$

To facilitate our discussion, we define the length function of the answer as follows. Since the length of the answer Y_t is determined by the query Q_t, let $\ell(Q_t)$ be the length of Y_t and the average length of the answer at time t is given by

$$\ell_t = \mathbb{E}_{Q_t}[\ell(Q_t)], \qquad (4)$$

where $\mathbb{E}[\cdot]$ is the expectation operator.

The query and answer functions need to satisfy the following decodable and privacy constraints.

1) Decodability: For any time t, the user should be able to recover the desired message from the answer with zero-error probability, *i.e.,*

$$H\left(W_{X_t,t} | Y_t\right) = 0, \quad \forall t \in \mathbb{N}. \qquad (5)$$

2) Privacy: For any time t, the user's requests over time where the privacy is required should not be revealed to the server, *i.e.,*

$$I\left(X_{\mathcal{B}_t}; Q_{[t]}\right) = 0, \quad \forall t \in \mathbb{N}, \qquad (6)$$

where $\mathcal{B}_t := \{i : i \leq t, F_i = \text{ON}\} \cup \{i : i \geq t+1\}$, and $[t] := \{0, 1, \ldots, t\}$.

We would like to clarify the privacy requirement in (6). The user does not know whether privacy is ON or OFF in the future. For this reason, we have adopted a worst-case formulation in the privacy constraint by assuming that privacy is always ON in the future.

For any message length L, the tuple $(\ell_t : t \in \mathbb{N})$ is said to be achievable if there exists a code satisfying the decodability and the privacy constraint. The efficiency of the code can be measured by L/ℓ_t. Hence, we define the achievable region by the convention as follows:

Definition 1. *The rate tuple $(R_t : t \in \mathbb{N})$ is achievable if there exists a code with message length L and average download cost ℓ_t such that $R_t \leq L/\ell_t$.*

Before proceeding to the results, we would like to mention that coded retrieval is not helpful in this problem. The point can be formally argued by dividing the possible queries to 2^N subsets, each of which corresponds to the decodability of a subset of the latest messages. Details can be found in [5]. For this reason, we only consider that Q_t takes value in $\mathcal{Q} = 2^{\mathcal{N}}$ in the following sections.

III. Main result

In this section, we present the main result of this paper, that is, the characterization of the achievable region for the two-sources system, *i.e.,* $N = 2$. For clarity, we will use A and B to denote the two sources, that is, each X_t takes values in $\mathcal{N} = \{A, B\}$. Correspondingly, the query Q_t takes values in $\mathcal{Q} = \{\{A\}, \{B\}, \{A, B\}\}$. We do not distinguish between A and $\{A\}$ in our notation, and $\{A, B\}$ will be written as AB.

Before stating our main result, we need to set up some useful notations. For simplicity, we assume that $F_0 = \text{ON}$. For any t, let $F^-(t) := \max\{i : i \leq t, F_i = \text{ON}\}$, *i.e.,* $F^-(t)$ is the latest time such that the privacy is ON. For our analysis, it is convenient to define $U_t := \left(X_{F^-(t)}, X_{t+1}\right) \in \mathcal{N}^2$, which represents the last request when privacy was ON and the next request of the user at time t.

We will need $p\left(x_t|u_t\right)$, which is given by

$$p\left(x_t|u_t\right) = \frac{p\left(x_{t+1}|x_t\right) p\left(x_t|x_{F^-(t)}\right)}{p\left(x_{t+1}|x_{F^-(t)}\right)}.$$

Here, $p\left(x_{t+1}|x_t\right)$, $p\left(x_t|x_{F^-(t)}\right)$ and $p\left(x_{t+1}|x_{F^-(t)}\right)$ can be determined from M, $M^{t-F^-(t)}$ and $M^{t+1-F^-(t)}$ respectively, where M is the transition matrix of the Markov chain representing the user's requests. Moreover, we introduce the following definition:

$$\pi(x_t) := \min_{u_t \in \mathcal{N}^2} p\left(x_t|u_t\right), \ \forall x_t \in \mathcal{N}. \tag{7}$$

In other words, if we write $p\left(x_t|u_t\right)$ as a $N^2 \times N$ probability transition matrix, $\pi(x_t)$ is the minimum value of each column.

Now, we are ready to state the main result in the following theorem.

Theorem 1. *Suppose that $\{X_t : t \geq 0\}$ is a Markov process with the transition matrix M. The rate tuple $(R_t : t \in \mathbb{N})$ is achievable if and only if*

$$\frac{1}{R_t} \geq 2 - \sum_{x_t \in \mathcal{N}} \pi(x_t). \tag{8}$$

To prove the theorem, we will give an explicit scheme that achieves the rate given in the R.H.S of (8) in Section IV and prove its optimality in Section V. Before that, we give an example to illustrate the rate given in (8).

Example 1. *Consider $(F_0, F_1) = (ON, OFF)$ and the transition matrix of the Markov chain is given by*

$$M = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}, \ 0 \leq \alpha \leq \frac{1}{2},$$

where $M_{i,j}$ is the transition probability from source i to source j (assuming source 1 is A and source 2 is B).

Consider the rate at $t = 1$. From (8), we have

$$\frac{1}{R_1} \geq 2 - \frac{2\alpha^2}{\alpha^2 + (1-\alpha)^2},$$

which means that it is not necessary for the user to download both messages except when $\alpha = 0$. When $\alpha = 0.5$, $R_1 \geq 1$. The reason is that at each time the user simply downloads only his/her desired message when the requests are independent.

Few remarks about the theorem are due here.

Remark 1. *In our model, we have assumed that the user knows the future requests within a window of positive size $\omega \geq 1$. An interesting implication of Theorem 1 is that the optimal rate does not depend on the window size. This means that increasing the window size into the future beyond one does not increase the rate. The case when the user does not know any future requests, i.e., $\omega = 0$, falls into a different model, which was studied in [5].*

Remark 2. *If $F_t = ON$, we have $U_t = (X_t, X_{t+1})$, and then we can easily see that R_t is achievable if and only if $R_t \leq \frac{1}{2}$ from (7) and (8), which means that it is necessary to download two messages. This is consistent with the well-known result [3].*

IV. Proof of Theorem 1: Achievability

A. ON-OFF Privacy Scheme

Here, we describe our query encoding function as defined in Section II. The query Q_t is encoded from X_t, U_t and S_t, *i.e.,*

$$Q_t = \phi_t\left(U_t, X_t, S_t\right).$$

Since we are not interested in the local randomness used, instead of writing ϕ_t explicitly, the function ϕ_t can be completely described by the probability distribution $w\left(q_t|x_t, u_t\right)$, which is given by

q_t	x_t	$\bar{x}_t$	AB
$w(q_t\|x_t,u_t)$	$\frac{\pi(x_t)}{p(x_t\|u_t)}$	0	$1-\frac{\pi(x_t)}{p(x_t\|u_t)}$

Here, $\bar{x}_t$ is defined as $\{A, B\}\backslash\{x_t\}$. Since $q_t \neq \bar{x}_t$ is always true, for notational simplicity, we write the encoding function $w(q_t|x_t, u_t)$ as

$$w\left(q_t|x_t, u_t\right) = \begin{cases} \frac{\pi(x_t)}{p(x_t|u_t)}, & |q_t| = 1, \\ 1 - \frac{\pi(x_t)}{p(x_t|u_t)}, & |q_t| = 2. \end{cases} \tag{9}$$

Example 2. *Let us adopt the same setting as in Example 1. Suppose that at time $t = 1$, the user wants source A, i.e., $X_1 = A$, and we need to determine the query Q_1. First, we determine*

$$\pi\left(x_1\right) = \frac{2\alpha^2}{1 + (1-2\alpha)^2}.$$

In our scheme in (9), Q_1 will be dependent on X_0 and X_2. Suppose that $X_0 = X_2 = A$, and then Q_1 will be given by

$$w\left(q_1|x_1, u_1\right) = \begin{cases} \frac{\alpha^2}{(1-\alpha)^2}, & |q_1| = 1, \\ \frac{1-2\alpha}{(1-\alpha)^2}, & |q_1| = 2. \end{cases}$$

In other words, if $X_0 = X_1 = X_2 = A$, then the user will toss a biased coin such that with probability $\frac{\alpha^2}{(1-\alpha)^2}$, he/she will download only the message generated by source A and with probability $\frac{1-2\alpha}{(1-\alpha)^2}$, he/she will download both messages.

B. Rate

We first show that the given coding scheme achieves the rate

$$R_t = \frac{1}{2 - \sum_{x_t} \pi(x_t)}.$$

Since

$$p\left(q_t\right) = \sum_{x_t, u_t} p\left(x_t, u_t\right) w\left(q_t|x_t, u_t\right),$$

by substituting (9), we have

$$p\left(q_t\right)=\begin{cases}\sum\limits_{x_t,u_t} p\left(u_t\right)\pi(x_t), & |q_t|=1,\\ 1-\sum\limits_{x_t,u_t} p\left(u_t\right)\pi(x_t), & |q_t|=2.\end{cases} \tag{10}$$

Note that $\pi(x_t)$ is independent of u_t, so (10) can be written as

$$p\left(q_t\right)=\begin{cases}\sum\limits_{x_t}\pi(x_t), & |q_t|=1,\\ 1-\sum\limits_{x_t}\pi(x_t), & |q_t|=2,\end{cases} \tag{11}$$

which immediately gives that

$$\frac{1}{R_t}=\frac{\ell_t}{L}=\mathbb{E}\left[|Q_t|\right]=2-\sum_{x_t}\pi(x_t).$$

C. Privacy

It remains to show that the encoding function given in (9) satisfies the privacy constraint in (6). We prove this by induction on t.

First, consider the base case where $t=0$. Since $F_0=\text{ON}$, we know that $Q_0=AB$ from (9), so we have

$$I\left(X_{\mathcal{B}_0};Q_{[0]}\right)=0.$$

Now, we start the inductive step. Assume that

$$I\left(X_{\mathcal{B}_{t-1}};Q_{[t-1]}\right)=0, \tag{12}$$

we need to show that

$$I\left(X_{\mathcal{B}_t};Q_{[t]}\right)=0.$$

Towards this end, consider

$$I\left(X_{\mathcal{B}_t};Q_{[t]}\right)=\underbrace{I\left(X_{\mathcal{B}_t};Q_{[t-1]}\right)}_{I_1}+I\left(X_{\mathcal{B}_t};Q_t|Q_{[t-1]}\right),$$

where $I\left(X_{\mathcal{B}_t};Q_t|Q_{[t-1]}\right)$, the second term in the summation above, can be written as

$$\begin{aligned}&I\left(X_{\mathcal{B}_t};Q_t|Q_{[t-1]}\right)\\ &= I\left(U_t;Q_t|Q_{[t-1]}\right)+I\left(X_{\mathcal{B}_t}\backslash U_t;Q_t|U_t,Q_{[t-1]}\right)\\ &= I\left(U_t;Q_{[t]}\right)-\underbrace{I\left(U_t;Q_{[t-1]}\right)}_{I_2}+\underbrace{I\left(X_{\mathcal{B}_t}\backslash U_t;Q_t|U_t,Q_{[t-1]}\right)}_{I_3}.\end{aligned}$$

Thus, we have

$$I\left(X_{\mathcal{B}_t};Q_{[t]}\right)=I\left(U_t;Q_{[t]}\right)+I_1-I_2+I_3. \tag{13}$$

Proposition 1. $I_1=I_2=I_3=0$.

This proposition is mainly due to the causality of the encoding function and the Markovity of the user's requests. The proof details will be given at the end of this section.

It remains to show that $I\left(U_t;Q_{[t]}\right)=0$, which can be equivalently written as $p\left(u_t|q_{[t]}\right)=p\left(u_t\right)$. To see this, consider

$$\begin{aligned}p\left(u_t|q_{[t]}\right)&=\sum_{x_t}p\left(u_t,x_t|q_t,q_{[t-1]}\right)\\ &=\sum_{x_t}\frac{p\left(u_t,x_t,q_t|q_{[t-1]}\right)}{p\left(q_t|q_{[t-1]}\right)}\\ &=\frac{\sum_{x_t}p\left(u_t,x_t|q_{[t-1]}\right)p\left(q_t|u_t,x_t,q_{[t-1]}\right)}{p\left(q_t|q_{[t-1]}\right)}\\ &=\frac{\sum_{x_t}p\left(u_t,x_t|q_{[t-1]}\right)p\left(q_t|u_t,x_t,q_{[t-1]}\right)}{\sum_{x_t,u_t}p\left(u_t,x_t|q_{[t-1]}\right)p\left(q_t|u_t,x_t,q_{[t-1]}\right)}\\ &\overset{\text{(a)}}{=}\frac{\sum_{x_t}p\left(u_t,x_t|q_{[t-1]}\right)w\left(q_t|u_t,x_t\right)}{\sum_{x_t,u_t}p\left(u_t,x_t|q_{[t-1]}\right)w\left(q_t|u_t,x_t\right)}\\ &\overset{\text{(b)}}{=}\frac{\sum_{x_t}p\left(u_t,x_t\right)w\left(q_t|u_t,x_t\right)}{\sum_{x_t,u_t}p\left(u_t,x_t\right)w\left(q_t|u_t,x_t\right)},\end{aligned} \tag{14}$$

where (a) follows because Q_t is a stochastic function of $\{U_t,X_t\}$ given in (9), and (b) follows because $\{u_t,x_t\}\subseteq\mathcal{B}_{t-1}$ and the inducative assumption (12).

From (9), we have

$$p\left(u_t,x_t\right)w\left(q_t|u_t,x_t\right)=\begin{cases}p\left(u_t\right)\pi(x_t), & |q_t|=1,\\ p\left(u_t,x_t\right)-p\left(u_t\right)\pi(x_t), & |q_t|=2.\end{cases}$$

For $|q_t|=1$, (14) can be written as

$$\begin{aligned}p\left(u_t|q_{[t]}\right)&=\frac{\sum_{x_t}p\left(u_t,x_t\right)w\left(q_t|u_t,x_t\right)}{\sum_{x_t,u_t}p\left(u_t,x_t\right)w\left(q_t|u_t,x_t\right)}\\ &=\frac{\sum_{x_t}p\left(u_t\right)\pi(x_t)}{\sum_{x_t,u_t}p\left(u_t\right)\pi(x_t)}\\ &\overset{\text{(a)}}{=}\frac{p\left(u_t\right)\sum_{x_t}\pi(x_t)}{\sum_{x_t}\pi(x_t)\sum_{u_t}p\left(u_t\right)}\\ &=p\left(u_t\right),\end{aligned} \tag{15}$$

where (a) follows because $\pi(x_t)$ is independent of u_t.

Similarly, we can also check for $|q_t|=2$ that

$$p\left(u_t|q_{[t]}\right)=p\left(u_t\right). \tag{16}$$

From (15) and (16), we can obtain that

$$I\left(U_t;Q_{[t]}\right)=0. \tag{17}$$

Therefore, by plugging (17) into (13) and using Proposition 1, we obtain

$$I\left(X_{\mathcal{B}_t};Q_{[t]}\right)=0,$$

which concludes our induction proof.

D. Proof of Proposition 1

First, we have

$$I_1=I\left(X_{\mathcal{B}_t};Q_{[t-1]}\right)\overset{\text{(a)}}{\leq}I\left(X_{\mathcal{B}_{t-1}};Q_{[t-1]}\right)\overset{\text{(b)}}{=}0,$$

where (a) follows because $\mathcal{B}_{t-1} = \mathcal{B}_t \cup \{t\}$ by definition, and (b) follows from the inductive assumption (12). Second,

$$I_2 = I\left(U_t; Q_{[t-1]}\right) \overset{\text{(a)}}{\leq} I\left(X_{\mathcal{B}_t}; Q_{[t-1]}\right) = I_1 \leq 0,$$

where (a) follows because $U_t \subseteq X_{\mathcal{B}_t}$ by definition.

Finally, we prove that $I_3 = 0$ as follows

$$\begin{aligned} I_3 &= I\left(X_{\mathcal{B}_t} \setminus U_t; Q_t | U_t, Q_{[t-1]}\right) \\ &\overset{\text{(a)}}{\leq} I\left(X_{\mathcal{B}_t} \setminus U_t; U_t, X_t, S_t | U_t, Q_{[t-1]}\right) \\ &\overset{\text{(b)}}{=} I\left(X_{\mathcal{B}_t} \backslash U_t; X_t | U_t, Q_{[t-1]}\right) \\ &= I\left(X_{\mathcal{B}_t} \setminus U_t; X_t | U_t\right) + I\left(X_{\mathcal{B}_t} \setminus U_t; Q_{[t-1]} | X_t, U_t\right) \\ &\quad - I\left(X_{\mathcal{B}_t} \setminus U_t; Q_{[t-1]} | U_t\right) \\ &\overset{\text{(c)}}{=} I\left(X_{\mathcal{B}_t} \setminus U_t; X_t | U_t\right), \end{aligned} \tag{18}$$

where (a) follows because Q_t is encoded from $\{U_t, X_t, S_t\}$, (b) follows because S_t is independent of $\{X_i : i \in \mathbb{N}\}$ and $Q_{[t-1]}$, and (c) can be justified because one can check that

$$I\left(X_{\mathcal{B}_t} \backslash U_t; Q_{[t-1]} | X_t, U_t\right) = I\left(X_{\mathcal{B}_t} \backslash U_t; Q_{[t-1]} | U_t\right) = 0$$

from $\mathcal{B}_{t-1} = \mathcal{B}_t \cup \{t\}$ and the inductive assumption (12).

To finish proving $I_3 = 0$, we claim that

$$I\left(X_{\mathcal{B}_t} \setminus U_t; X_t | U_t\right) = 0.$$

Towards this end, by letting $\mathcal{B}_t^- = \{i : i \leq t, F_i = \text{ON}\} \setminus \{F^-(t)\}$, and $\mathcal{B}_t^+ = \{i : i \geq t+2\}$, we can easily obtain from the Markovity of $\{X_i : i \in \mathbb{N}\}$ that

$$I\left(X_{\mathcal{B}_t} \backslash U_t; X_t | U_t\right) = I\left(X_{\mathcal{B}_t^-}, X_{\mathcal{B}_t^+}; X_t | X_{\{F^-(t), t+1\}}\right) = 0,$$

which concludes that $I_3 = 0$.

V. Proof of Theorem 1: Converse

To obtain an upper bound on the rate R_t, we derive a lower bound on the average downloading cost $\mathbb{E}\left[|Q_t|\right]$, which can be obtained by solving the following optimization problem:

$$\begin{aligned} \underset{p(u_t, x_t, q_t)}{\text{minimize}} \quad & \mathbb{E}\left[|Q_t|\right] = \sum_{q_t} p\left(q_t\right) |q_t| \\ \text{subject to} \quad & p\left(x_t, q_t\right) = 0, \ x_t \notin q_t, \qquad \text{(decodability)} \\ & p\left(q_t | u_t\right) = p\left(q_t\right). \qquad \text{(relaxed privacy)} \end{aligned} \tag{19}$$

Here, the relaxed privacy constraint is obtained by relaxing our original privacy requirement $I\left(Q_{[t]}; X_{\mathcal{B}_t}\right) = 0$ to $I\left(Q_t; U_t\right) = 0$. This is a relaxation because $\{F^-(t), t+1\} \subseteq \mathcal{B}_t$.

For clarity, we illustrate all feasible $p\left(u_t, x_t, q_t\right)$ in Table I with two auxiliary variables z_1 and z_2, where

$$z_1 = \Pr\left(Q_t = A | U_t = (A, A)\right),$$

and

$$z_2 = \Pr\left(Q_t = B | U_t = (A, A)\right).$$

Clearly, all entries in Table I must be non-negative.

U_t	X_t	$Q_t = A$	$Q_t = B$	$Q_t = \{A, B\}$
(A, A)	A	$z_1 p_{aa}$	0	$p_{aa}\left(p_{a\|aa} - z_1\right)$
(A, A)	B	0	$z_2 p_{aa}$	$p_{aa}\left(p_{b\|aa} - z_2\right)$
(A, B)	A	$z_1 p_{ab}$	0	$p_{ab}\left(p_{a\|ab} - z_1\right)$
(A, B)	B	0	$z_2 p_{ab}$	$p_{ab}\left(p_{b\|ab} - z_2\right)$
(B, A)	A	$z_1 p_{ba}$	0	$p_{ba}\left(p_{a\|ba} - z_1\right)$
(B, A)	B	0	$z_2 p_{ba}$	$p_{ba}\left(p_{b\|ba} - z_2\right)$
(B, B)	A	$z_1 p_{bb}$	0	$p_{bb}\left(p_{a\|bb} - z_1\right)$
(B, B)	B	0	$z_2 p_{bb}$	$p_{bb}\left(p_{b\|bb} - z_2\right)$

TABLE I: The joint distribution $p\left(u_t, x_t, q_t\right)$ satisfying the decodability and the privacy constraint, where p_{aa} denotes $\Pr\left(U_t = (A, A)\right)$, and $p_{a|aa}$ denotes $\Pr\left(X_t = A | U_t = (A, A)\right)$. Both are constants given by the transition matrix of the Markov chain.

Then the optimization problem given in (19) can be rewritten as

$$\begin{aligned} \underset{z_1, z_2}{\text{minimize}} \quad & \mathbb{E}\left[|Q_t|\right] = 2 - z_1 - z_2 \\ \text{subject to} \quad & 0 \leq z_1 \leq \pi(A), \\ & 0 \leq z_2 \leq \pi(B), \end{aligned} \tag{20}$$

where $\pi(A)$ and $\pi(B)$ are defined in (7).

We can easily see that the optimal value to the problem in (20) is given by

$$\min_{z_1, z_2} (2 - z_1 - z_2) = 2 - \pi(A) - \pi(B),$$

which completes the proof that

$$\frac{1}{R_t} \geq 2 - \sum_{x_t \in \{A, B\}} \pi(x_t).$$

References

[1] L. Sweeney, "K-anonymity: A model for protecting privacy," *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, vol. 10, no. 5, pp. 557–570, Oct. 2002.

[2] C. Dwork, "Differential privacy," in *33rd International Colloquium on Automata, Languages and Programming (ICALP)*, 2006.

[3] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *IEEE Symposium on Foundations of Computer Science*, 1995.

[4] P. Dhungel, M. Steiner, I. Rimac, V. Hilt and K. W. Ross, "Waiting for Anonymity: Understanding Delays in the Tor Overlay," in *IEEE Tenth International Conference on Peer-to-Peer Computing (P2P)*, Delft, 2010.

[5] C. Naim, F. Ye, and S. El Rouayheb, "ON-OFF privacy with correlated requests," *arXiv:1905.00146*, 2019.

[6] N. Shah, K. Rashmi, and K. Ramchandran. "One extra bit of download ensures perfectly private information retrieval," in *IEEE International Symposium on Information Theory (ISIT)*, 2014.

[7] H. Sun and S. Jafar, "The capacity of private information retrieval," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 40754088, 2017.

[8] K. Banawan and S. Ulukus, " The capacity of private information retrieval from coded databases," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1945-1956, 2018.

[9] R. Tajeddine and S. El Rouayheb, "Private information retrieval from mds coded data in distributed storage systems," in *IEEE International Symposium on Information Theory (ISIT)*, 2016.

[10] S. Kadhe, B. Garcia, A. Heidarzadeh, S. El Rouayheb, and A. Sprintson, "Private information retrieval with side information: The single server case," in *55th Annual Allerton Conference on Communication, Control, and Computing*, 2017.